

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 569)
 AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 JOURNAL Tumor Gene Index
 COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: ccapbs@mail.nih.gov
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
 NCI-CCAP clone distribution information can be found through the
 I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html The following repetitive
 elements were found in this cDNA sequence: 1-22,
 >Al-rich/Low-complexity
 Seq primer: M13 Forward
 POLYA=yes.

FEATURES

source
 1..569
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3087997"
 /clone_lib="NCI-CCAP_Sub5"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not 1; Site 2: Eco RI; NCI-CCAP Sub5
 is a subtracted library derived from NCI-CCAP_Sub5. The
 NCI-CCAP_Sub8 library had 2.5 million recombinants. A
 single-stranded DNA preparation of NCI-CCAP_Sub5 was used
 as a tracer in a subtractive hybridization with a driver
 comprising a pool of clones from NCI-CCAP_Sub5 (IMAGE
 clone ids 2732833-2737415, 3068040-3069191; 25% of the
 driver population), a pool of clones from NCI-CCAP_Sub4
 (IMAGE clone ids 2723592-2729326; 25% of the driver
 population), NCI-CCAP_Sub5 (pool AIF-AJU, IMAGE ids
 2728959-273190; 25% of the driver population), and
 NCI-CCAP_Sub7 (IMAGE ids 3069192-3072238, 3081854-3084550
 : 25% of the driver population). Subtraction was
 performed as previously described [Bonaldo, Lennon &
 Soares (1996): Normalization and Subtraction: Two
 Approaches To Facilitate Gene Discovery. Genome Research
 6, 791-806.
 TAG_LIB=NCI-CCAP_GC4
 TAG_TISSUE=germ cell
 TAG_SEQ=AAATC"
 BASE COUNT 149 a 142 c 95 g 183 t
 OPIGIN

Query Match 85.2% Score 397; DB 10; Length 569;
 Fast Local Similarity 98.0%; Pred. No. 3.2e-105;
 Matches 444; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

OY 1 atcaagaacataaagttcggacaataattcttcctaccatcccaacaaatctta 60
 |||||||
 Db 95 ATCAAGAAACATAGAGTII-GGGCAATATACITTCATCCATCCACCCCAATCTTA 153
 OY 61 ctctactcattctctctcattatttgggaataatcatcagaagatgtgtcttaagta 120
 |||||||
 Db 154 CTCTACTCATCTCATCTCATTAATTTGGGAATATCATGAGAGATGTGTCTTGAGTA 213
 OY 121 agagattaaagaataaagctttttgacccctgccacaccccccatgccaggagtgatcac 180
 |||||||
 Db 214 AGAGATTAAAGAAATTAAGCTTTTTCACCCCTGCCAACACCCCATCCCGGGTGTGTCAC 273
 OY 181 cctccatacaataacatccaggaagaataagtctccctttctcgaatccgtaattcgcc 240
 |||||||
 Db 274 CTCCCAATACATAGATGCCGGAAGAGTAGTTCCTCTTCTGTAGTCGGTAATCTGCC 333

OY 241 atcatcttcccatcttccagttctctcttccattgcaagtccacaatctggtctcaggat 300
 |||||||
 Db 334 ATCATCTTCCATCTTCCAGTCT-CTTTCATTCGAAGTCACATCTGGGTCTCAGGAT 392
 OY 391 tataccgctattatctcgatctattgtcttctcacttctgcccactgagctgacattcgac 360
 |||||||
 Db 393 TATACCGCTTATAGTCTCGATCATCTTTCACCTTGTCACCTGAGCTGACCTTCGCAC 452
 OY 361 ctggagagagtcctcttcttcctcatcactcactcaccacaaagaaagagcagctcct 420
 |||||||
 Db 453 CT-GGAGGAGTGCTCTTTCCTTCATCACCCTGACCTCCACAGAAACAGGGCAGCTCT 511
 OY 421 catcatgggcttccaccactttcaggggtaag 453
 |||||||
 Db 512 CATCACT-GGCTTCACCACTTTCAGGTAAGG 543
 RESULT
 BF511148
 LOCUS
 DEFINITION
 BF511148
 ACCESSION
 VERSION
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 571)
 AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 JOURNAL Tumor Gene Index
 COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: ccapbs@mail.nih.gov
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. cDNA library Preparation: M.B. Soares Lab Clone distributi
 NCI-CCAP clone distribution information can be found through the
 I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html The following repetitive
 elements were found in this cDNA sequence: 1-22,
 >Al-rich/Low-complexity
 Seq primer: M13 Forward
 POLYA=yes.

FEATURES
 Location/Qualifiers
 1..571
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3087989"
 /clone_lib="NCI-CCAP_Sub8"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not 1; Site 2: Eco RI; NCI-CCAP_Sub5
 is a subtracted library derived from NCI-CCAP_Sub5. The
 NCI-CCAP_Sub8 library had 2.5 million recombinants. A
 single-stranded DNA preparation of NCI-CCAP_Sub5 was used
 as a tracer in a subtractive hybridization with a driver
 comprising a pool of clones from NCI-CCAP_Sub5 (IMAGE
 clone ids 2732833-2737415, 3068040-3069191; 25% of the
 driver population), a pool of clones from NCI-CCAP_Sub1
 (IMAGE clone ids 2723592-2729326; 25% of the driver
 population), NCI-CCAP_Sub6 (pool AIF-AJU, IMAGE ids
 2728959-273190; 25% of the driver population), and
 NCI-CCAP_Sub7 (IMAGE ids 3069192-3072238, 3081854-3084550
 : 25% of the driver population). Subtraction was
 performed as previously described [Bonaldo, Lennon &
 Soares (1996): Normalization and Subtraction: Two
 Approaches To Facilitate Gene Discovery. Genome Research
 6, 791-806.
 TAG_LIB=NCI-CCAP_GC4

[illegible]

Db	441	CT-GGAGGAGTGCTCTGTGCTCTATCATCTGACTGCACAAAGAAACAGGCGAGCTCT	499
Qy	421	catcacggagcttcaccactttcaggagtaag	453
Db	500	CATCAGT-GGCTTCACCACTTTCAGSGTAAGG	531
RESULT	7		
LOCUS	AI656143	750 bp mRNA linear	EST 04-MAY-1999
DEFINITION	tt43f03.x1 NCI-CGAP GC6 Homo sapiens cDNA clone IMAGE:2243549		3
ACCESSION	AI656143	similar to TR:015205 OL5205 DIUBIQUITIN. ?; mRNA sequence.	
VERSION	AI656143.1	GI:4740122	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 750)		
FEATURES	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
COMMENT	National Cancer Institute. Cancer Genome Anatomy Project (CGAP).		
	Tumor Gene Index		
	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaps-r@mail.nih.gov		
	Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael		
	R. Emmert-Buck, M.D., Ph.D.		
	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima		
	Bonaldo, Ph.D.		
	cDNA Library Arrayed by: Greg Lennon, Ph.D.		
	DNA Sequencing by: Washington University Genome Sequencing Center		
	Clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. consortium/LNL at:		
	www-bio.llnl.gov/hbrp/image/image.html		
	Seq primer: -400p from Gibco		
	High quality sequence stop: 450.		
FEATURES	Location/Qualifiers		
SOURCE	1..750		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:2243549"		
	/clone_lib="NCI-CGAP_GC6"		
	/tissue_type="pooled germ cell tumors"		
	/lab_host="DH10b"		
	/note="Vector: pI73D-Pac (Pharmacia) with a modified		
	polylinker: Site_1: Not 1; Site_2: Eco RI; Plasmid DNA		
	ss circles were made in vitro. Following HAP purification, and		
	this DNA was used as tracer in a subtractive hybridization,		
	reaction. The driver was PCR-amplified cDNAs from a pool		
	of 5,000 clones made from the same library (cloneIDs		
	1257096-1258631, 1459064-1470983, and 1475592-1476743).		
	Subtraction by Bento Soares and M. Fatima Bonaldo.		
BASE COUNT	197 a 186 c 142 g 213 t	10 others	
ORIGIN			
	Query Match	85.0%	Score 396; DB 9; Length 750;
	Best local Similarity	97.8%	Pred. No. 6.8e-105;
	Matches	443; Conservative	0; Mismatches 6; Indels 4; Gaps 4;
Qy	1	atcaagaacatagadttcggcgcaatatacttcattccatccaccacccaatactta	50
Db	03	ATCAAGAAACATAGAGTT-GGGCAATATATCTTCATCTACCCATCCACGCCAATCTTA	141
Qy	61	ctctactcattctcatttaatttggaaatcatcagaagatgtgttcattgaagta	120
Db	142	CTCTACTATCTCATCTCAITANTTTTGGGAATATCAGAGATGTGTTCGTTAGTA	201
Qy	101	agagattaaagaataagctttttagaccctgcacaacccccatgccaggatggtcac	180
Db	242	AGAGATTAAAGAAATTAAGTTTITGACCTCGCAACACCCATCCAGGGTGGTTCAC	261

NCL_CGAP_Sub8 library had 2.5 million recombinants. A single-stranded DNA preparation of NCL_CGAP_Sub5 was used as a tracer in a subtractive hybridization with a driver comprising a pool of clones from NCL_CGAP_Sub5 (IMAGE clone ids 2732833-2737413, 3056040-3069233; 25% of the driver population), a pool of clones from NCL_CGAP_Sub4 (IMAGE clone ids 2723552-2729336; 25% of the driver population), NCL_CGAP_Sub6 (pool AIF-AJL; IMAGE ids 2728959-2733190; 25% of the driver population), and NCL_CGAP_Sub7 (IMAGE ids 3069162-3072229, 3081863-3084550; 25% of the driver population). Subtraction was performed as previously described [Bonaldi, Lennon & Soares (1995) Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery, Genome Research 5, 791-806].

TAG_LIB=NCI_CGAP_GC4
TAG_TISSUE=germ cell
TAG_SFO=AAATC

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
CONTACT Contact: Robert Stancovski

/tissue_type="well-differentiated endometrial;
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014."

BASE COUNT 166 a 160 c 119 g 200 t
ORIGIN

Query Match 82.8%; Score 385.8; DB 9; Length 545;
Best Local Similarity 96.5%; Pred. No. 6.2e-102;
Matches 437; Conservative 0; Mismatches 12; Indels 4; Gaps 4:
Qy 1 atcaagaacacatagattcgggcaataatactctacccatccacccacaaatttta 60
|||||
Db 96 ATCAAGAACAATAGATT-GGGCAATATCTTCAICCTACCCATCCCAACAACTTTA 154
Qy 61 ctctact 120
|||||
Db 155 CTCTACT 214
Qy 121 agagattataaataaataaagctttttgacccctgccaacaccccatgcccagagctttac 180
|||||
Db 215 AGAGATTATAAAGAAATAAGCTTTTGGACCCCTGCCAACACCCCATCCCGAGGIGHCAC 274
Qy 181 ctctcaatacaatacaatcaccgaagaagtaagttgacctctctctctctctctctct 240
|||||
Db 275 CTCTCAATAACAATAAGATGCCAGGAAGTAAGTTTGGCCCTTCTGATCGGTAATCTGCC 334
Qy 241 atcatcttcccatcttccagctctctctctctctctctctctctctctctctctct 300
|||||
Db 335 ATCATCTTCCCATCTTCCAGTCT-CTTTGATTCGAAGTCACAATCTGGGCTCAGCAT 393
Qy 301 tataccgctctagttctcatctgctctctctctctctctctctctctctctctctctct 360
|||||
Db 394 TATACCGCTCTTAGTCTCGATCATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 453
Qy 361 ctggagagagatgacctcttctgctcctcatcctgactccacagaacacgaagggcagctct 420
|||||
Db 454 CT-GGAGGAGGTGCCCTCTTCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 512
Qy 421 catcactggggcttccaccacttccaggggtaag 453
|||||
Db 513 CATCACT-GGGCTTCACCACTTTCAGGTAAGG 544

RESULT 15
AI985834
LOCUS
DEFINITION
w78c12.x1 NCI-CCAP_Utl1 Homo sapiens cDNA clone IMAGE:2493814 3'
similar to TR:015205 O15205 DIUBIQUITIN..; mRNA sequence.
ACCESSION
AI985834
VERSION
AI985834.1 GI:5813111
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 501)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapb-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D. Michael E.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
***bio.llnl.gov/abrp/image/image.html
Insert length: 678 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 421.
Location/Qualifiers

FEATURES
Source

1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI-CCAP_Utl1"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014."
BASE COUNT 150 a 154 c 107 g 179 t 1 others
ORIGIN

Query Match 82.6%; Score 385; DB 9; Length 501;
Best Local Similarity 97.6%; Pred. No. 1e-101;
Matches 443; Conservative 0; Mismatches 6; Indels 5; Gaps 5:
Qy 1 atcaagaacacatagattcgggcaataatactctacccatccacccacaaatttta 60
|||||
Db 78 ATCAAGAACAATAGATT-GGGCAATATCTTCAICCTACCCATCCCAACAACTTTA 136
Qy 61 ctctact 120
|||||
Db 137 CTCTACT 196
Qy 121 gaaataaagaataaagctttttgacccct-gccaacaccccatgcccaggtggtca 179
|||||
Db 137 AGAGATTATAAAGAAATAAGCTTTTGGACCCCTGCCAACACCCCATCCCGAGGIGTCA 256
Qy 150 cctctcaatacaataacatgcccaggaagtaagttgaccttctctgagcgtaactgac 239
|||||
Db 257 CCTCTCAAATACAAATAGATGCCAGGAAGTAAGTTGCCCTTCTGATGCCGTAATCTGC 316
Qy 241 tataccgctctagttctcatctgctctctctctctctctctctctctctctctctctct 299
|||||
Db 317 TATACCGCTCTTAGTCTCGATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 375
Qy 350 ttataccgctctagttctcatctgctctctctctctctctctctctctctctctctctct 359
|||||
Db 376 TATACCGCTCTTAGTCTCGATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 435
Qy 350 cctggagagagatgaccttctgctcctcatcactgactccacagaacacgaagggcagctcc 419
|||||
Db 436 CCGT-GGAGGAGGTGCCCTCTNIGCCCTCATCCTGACTCTCACAAGAACAAGGCGAGCTCC 494
Qy 420 tcatacactggggcttccaccacttccaggggtaag 453
|||||
Db 495 TCATCACT-GGGCTTCACCACTTTCAGGTAAGG 527

Search completed: July 29, 2002, 17:20:22
Job time: 7743 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 17:55:00 : Search time 1928.76 seconds
(without alignments)
3992.703 Million cell updates/sec

Title: US-09-880-107-3847

Perfect score: 368
Sequence: 1 gaagagacgtgtaagtgcg.....ataaatgtatctctatc 369

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 359512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_em.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_un.*

28: em_vl.*

29: em_hg.*

30: em_hg_hum.*

31: em_hg_inv.*

32: em_hg_inv.*

33: em_hg_inv.*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

Result No. Score Match Length DB ID Description

1	368	100.0	368	6	AX332609	Sequence
2	368	100.0	368	6	AX333248	Sequence
3	368	100.0	368	9	HSPST1	Y00705 Homo sapien
4	344.8	93.7	608	6	AX014216	Sequence
5	340.8	92.6	432	9	HUMPS11	M11949 Human pancr
6	307.8	83.6	341	6	E01574	CDNA sequen
7	211.4	57.4	265	6	AR166862	Sequence
8	211.4	57.4	265	6	AX192488	Sequence
9	168.8	45.9	179	6	E01725	Synthetic D
10	168.8	45.9	179	6	E02456	DNA encodin
11	162.2	44.2	341	12	SYNECOPST	M22196 Synthetic E
12	162.2	44.1	420	6	AR166863	Sequence
13	162.2	44.1	420	6	AR192489	Sequence
14	162	44.0	302	6	AR059723	Sequence
15	162	44.0	302	6	E08411	DNA sequenc
16	160.8	43.7	234	6	E01819	Sequence
17	160.8	43.7	234	6	E03120	DNA sequenc
18	160	43.5	204	12	SYNPSTIB	M11103 Synthetic h
19	159.6	43.4	379	10	RATPSTIB	M27883 Rat pancrea
20	158	42.9	377	10	RNMNPEP	X59696 R.norvegicu
21	156.4	42.5	359	10	RATPSTIA	M27882 Rat pancrea
22	145.8	39.9	369	10	MMP1MP12	X06342 Mouse mRNA
23	144.6	39.3	238	10	RATPSTIAA	M35299 Rat pancrea
24	139.8	38.0	317	10	RATMONPA	M22162 Rat monitor
25	119	32.3	422	9	HUMPS1A1	M2971 Human pancr
26	119	32.3	7370	9	AF286028	Homo sapi
27	119	32.3	150255	2	AC011352	Homo sapi
28	119	32.3	165889	2	AC011402	Homo sapi
29	105.8	29.0	127	9	HUMPS1A3	M20529 Human pancr
30	103.8	28.2	182	6	A11350	Artificial
31	103.8	28.2	182	6	A11352	Artificial
32	103.8	28.2	182	6	A24905	PS11 master
33	103.8	28.2	182	12	SYNPSTIAA	M21832 Synthetic h
34	103.8	28.2	185	6	A24907	PS11 master
35	102.6	27.9	482	6	E03479	CDNA encodi
36	96.4	26.2	289	6	A14318	cl131-PS11-
37	96.4	26.2	289	6	A14319	cl131-PS11-
38	96.4	26.2	289	6	I08731	Sequence 19
39	96	26.1	168	6	I08731	Sequence 19
40	95	25.8	226	9	HUMPS1A4	M20530 Human pancr
41	89.2	24.2	150123	2	AC011344	Homo sapi
42	80.6	21.9	7008	6	AX251285	Sequence
43	73.2	19.9	7008	6	AX251286	Sequence
44	43.8	17.3	147	10	RATPSTI3	D11324 Rattus norv
45	33.6	17.3	54541	2	AC104867	Mus muscu

ALIGNMENTS

RESULT 1
AX332609
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
score
BASE C N1 100 a 79 c 89 g 100 t

AX332609 368 bp DNA linear PAT 09-JAN-2002

Sequence 3118 from Patent WO0194629.

AX332609

AX332609.1 GL:18123243

Human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (sites)

Young.P.E., Augustus.M., Carter.K.C., Ebner.R., Endress.G.,

Horrigan.S., Soppet.D.R. and Weaver,Z.

Cancer gene determination and therapeutic screening using signature

gene sets

Patent: WO 0194629-A 3118 13-DEC-2001;

Avalon Pharmaceuticals (US)

Location/Qualifiers

1..368

/organism="Homo sapiens"

/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 368; DB 6; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.4e-108;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gaagagacgtgtaagtgcggtgcggttttcaactgaacctctggagcagagaacttcagcc 60
DB 1 GAAGAGACGTGTAGTGCAGTGCAGTTCACACGACCTCTGGAGCGAGACCTTCAGCC 60
OY 61 atgaaggttaacaggcatctttcttcagtcgcttggccctcttgaagtcatactgataac 120
DB 61 ATGAAGGTAAACAGGCACTCTTCTCAGTCGCTTGGCCCTCTGAGCTACTACGCTAAC 120
OY 121 actgagctgactccctcgggaagagagagcacaatttcaaatgaacttaactgaatgcacc 180
DB 121 ACTGGAGCTGACTCCCTCGGAAGAGAGGCGCAAAATGTATCAATGAACCTTAATGSAATGCACC 180
OY 181 aagatatatgacctgtctgtgggactgaatgaataacttcccaatgaatgcctgatta 240
DB 181 AAGATATATGACCCCTGCTGTGGGACTGATGGAATACTTATCCCAATGAATGCGCTGTTA 240
OY 241 ttttttgaaggtcggaagacgcagactctctctcattcaaaaactcgggacttcgctga 300
DB 241 TGTITTTGAAGGTTCGGAAGCGCAGACCTTCTATCTCTCATTCATCAAAAATCGGGCTTGCCTGA 300
OY 301 gaaccaaggttttgaataatcccatcaggtcaccgcgagagccctattttgaataaaatgtatc 360
DB 301 GAACCAAGGTTTGAATAATCCCATCAGGTACAGTCACCGCGAGGCCATTGTTGAATAAATGATC 360
OY 361 tgaatc 368
DB 361 TGAATATC 368

RESULT 2

AX333248 AX333248 368 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 3757 from Patent WO0194629.
DEFINITION AX333248
ACCESSION AX333248
VERSION AX333248.1 GI:18123882
KEYWORDS
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrikan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets

JOURNAL

Patent: WO 0194629-A 3757 13-DEC-2001;

FEATURES

Avalon Pharmaceuticals (US)

Location/Qualifiers

1..368

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 100 a 75 c 89 g 100 t

ORIGIN

Query Match 100.0%; Score 368; DB 6; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.4e-108;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gaagagacgtgtaagtgcggtgcggttttcaactgaacctctggagcagagaacttcagcc 60
DB 1 GAAGAGACGTGTAGTGCAGTGCAGTTCACACGACCTCTGGAGCGAGACCTTCAGCC 60
OY 61 atgaaggttaacaggcatctttcttcagtcgcttggccctcttgaagtcatactgataac 120
DB 61 ATGAAGGTAAACAGGCACTCTTCTCAGTCGCTTGGCCCTCTGAGCTACTACGCTAAC 120

OY 121 actgagctgactccctcgggaagagagcacaatttcaaatgaacttaactgaatgcacc 180
DB 121 ACTGAGCTGACTCCCTCGGAAGAGAGGCGCAAAATGTATCAATGAACCTTAATGSAATGCACC 180
OY 181 aagatatatgacctgtctgtgggactgaatgaataacttcccaatgaatgcctgatta 240
DB 181 AAGATATATGACCCCTGCTGTGGGACTGATGGAATACTTATCCCAATGAATGCGCTGTTA 240
OY 241 ttttttgaaggtcggaagacgcagactctctcattcaaaaactcgggacttcgctga 300
DB 241 TGTITTTGAAGGTTCGGAAGCGCAGACCTTCTATCTCTCATTCATCAAAAATCGGGCTTGCCTGA 300
OY 301 gaaccaaggttttgaataatcccatcaggtcaccgcgagagccctattttgaataaaatgtatc 360
DB 301 GAACCAAGGTTTGAATAATCCCATCAGGTACAGTCACCGCGAGGCCATTGTTGAATAAATGATC 360
OY 361 tgaatc 368
DB 361 TGAATATC 368

RESULT 3

AX333248 AX333248 368 bp mRNA linear PRI 12-SEP-1993
LOCUS HSPST1
DEFINITION HSPST1
ACCESSION Y00705
VERSION Y00705.1 GI:35765
KEYWORDS psi1 gene; trypsin inhibitor.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 368)

Tomita,N.

Direct Submission

Submitted (15-NOV-1987) Tomita N., Institute of Molecular and

Cellular Biology, Osaka University, Yamada-oka, Suita 565, Japan

2 (bases 1 to 368)

Tomita,N., Horii,A., Yamamoto,T. and Ootawa,M.

Expression of pancreatic secretory trypsin inhibitor (PSTI) gene in

neoplastic tissues

FEBS Lett. (1987) In press

This sequence is identical with that of PSTI RNA <M11949> isolated

from human pancreatic cDNA library.

Location/Qualifiers

1..368

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="lambda TI-CI"

/tissue_type="sigmoid colon"

51..300

/note="inhibitor (AA 1-79)"

/codon_start=1

/protein_id="CAA68697.1"

/db_xref="GI:35765"

/db_xref="SNJSS-PROT:P00995"

/translation= MKVIGIFLLSALLSLSGNTGADSLGKAKYCNELNGCTKIYD

FVCGIDGNTYPECVLCFEGKRRKTSILIQKSGPC

349..354

/note="polyA signal"

368

/note="polyA site"

BASE COUNT 100 a 79 c 89 g 100 t

ORIGIN

Query Match 100.0%; Score 368; DB 9; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.4e-108;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 aaagacgtgtaagtgcggtgcggttttcaactgaacctctggagcagagaacttcagcc 60


```

QY 303 accaaggttttgaatcccatcaggtcaccgcgagcc-----tatgttgaata 352
|||||
DB 359 ACCAAGGTTTGAATCCCATCAGGTACCGCGAGGCTGACTGGCCITATGTTGAATA 418
QY 353 aatgtatctgaata 356
|||||
DB 419 AATGTAATCTGAATA 432

RESULT 6
LOCUS E01574 341 bp RNA linear PAT 29-SEP-1997
DEFINITION CDNA sequence of human PST1.
ACCESSION E01574
VERSION E01574.1 GI:2169827
KEYWORDS JP 1988098397-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 341)
AUTHORS Ogawa,M. and Matsubara,K.
TITLE PRODUCTION OF HUMAN PST1 BY YEAST
JOURNAL Patent: JP 1988098397-A 1 28-APR-1989;
SHIONOGI & CO LTD
COMMENT OS human
PN JP 1988098397-A/1
PD 28-APR-1988
PF 14-OCT-1985 JP 1985245049
PI OGAWA MICHIO, MATSUBARA KENICHI
PC C12P21/02.C12N15/00//C12N9/99.(C12P21/02.C12R1:645): CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=pancreas;
FH Key Location/Qualifiers
FT 5'UTR 1..25
FT CDS 26..265
FT 3'UTR
FT 256..341
FT sig_peptide 26..94
FT mat_peptide 95..262
FT /product='human PST1'
FT Location/Qualifiers
FT 1..341
FT /organism='Homo sapiens'
FT /db_xref='taxon:9606'
BASE COUNT 94 a 75 c 78 g 93 t
ORIGIN

Query Match 83.6%; Score 307.8; DB 6; Length 341;
Best Local Similarity 95.5%; Pred. No. 5.3e-89;
Matches 329; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 36 gactcttgagtcattcagccatcagccatgaaggtaacagggatctttcttcagtgccct 95
|||||
DB 1 GACCTCTGGAGCGAGAACITTCAGGCAITCAAGGTAAACAGGCAITCTTCTCAGTGGCTT 60
QY 96 gaccctgttgagttctatctgttaacactgaagctgactccctgggaagagagggccaaatg 155
|||||
DB 61 GGCCCTGTGAGTCTATCTGTTACACTGGAGCTGACTCCCTGGGAAGAGAGGCCAAATG 120
QY 156 ttacaatgaactaatggaatcaccagaatataatgacctgtctgtgtggaactgaaaga 215
|||||
DB 121 TTACAATGAACCTTAATGGATGCACCAAGATATATGACCCCTCTCTGTGGGAGTGAAGAA 180
QY 216 tacttatcccaatgaatcgtgttatcttttgaaggtcgaaacgccagacattctatcct 275
|||||
DB 181 TACTTATCCCAATGAATGCTGTGTTAIGTTTTTGAATAATCGGAAACGCCAGACTTCACTT 240

```

```

QY 276 catlcaaaaatctgggcttctcagaaaccaaagttttgaaatcccatcaggltaccgcg 335
|||||
DB 241 CATTCAAAATCTGGGCTTCTCTGAGAACCAAGGTTTTTGAATCCCATCAGTCAACCGCG 300
QY 336 aggc-----tatgttgaataaatgtatctgaata 366
|||||
DB 301 AGGCTGACACGGCCITATGTTGAATAATGATATCTGAATA 341

RESULT 7
LOCUS ARI65862 265 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 55 from patent US 6284241.
ACCESSION ARI65862
VERSION ARI65862.1 GI:16243230
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 265)
AUTHORS Xu,J.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
JOURNAL Patent: US 6284241-A 55 04-SEP-2001;
FEATURES Location/Qualifiers
FEATURES 1..265
source /organism='unknown'
BASE COUNT 69 a 59 c 57 g 75 t 5 others
ORIGIN

Query Match 57.4%; Score 211.4; DB 6; Length 265;
Best Local Similarity 95.6%; Pred. No. 1.1e-57;
Matches 237; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 76 atctttctcagtcacctgacctgttgagctctatctgttaacacacgtgagctaacctc 135
|||||
DB 1 ATCTTCTCTCAGTGGCTTGGCCNCTTGGAGTCTATCTGTTAACACACGTGAGCTGACCTC 60
QY 136 ctgggaagagagggccaaatgttacaatgaacttaataatgaatgacacacaaatatataacct 195
|||||
DB 61 CTGGGAAGAGAGGCCAAATGTTACAATGAACCTTAATGGATGCACCAAGATATATGACCT 120
QY 196 aactatgagactgaggaataacttaccatcccaatgaatg-cgtgttatg-ttttgaagctc 253
|||||
DB 121 GTCTGTTGGGACGTGTTGGAATACTTATCCCAATGAATGCCGTGTTATGTTTTTGAATAATC 180
QY 254 gaaaacggcagactctctatcctcattcgaataatcgggacctgtgtgagaacaaagtttt 313
|||||
DB 181 GGAACGCCAGACTTCTATCTCTCAITCAAAAATCTGGGCCCTTCTGAAACCCAGGGTTT 240
QY 314 gaaatccc 321
|||||
DB 241 NAAATCC 248

RESULT 8
LOCUS AX192488 265 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 55 from Patent WO0149716.
ACCESSION AX192488
VERSION AX192488.1 GI:15210452
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 265)
AUTHORS Xu,J., Lodes,M.J., Secret,H., Benson,D.R., Meagher,M.J.,
Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
JOURNAL Patent: WO 0149716-A 55 12-JUL-2001;

```


CORIXA CORPORATION (US)									
Location/Qualifiers									
1..265									
/organism="Homo sapiens"									
/db_xref="taxon:9606"									
BASE COUNT 69 a 59 c 57 g 75 t 5 others									
ORIGIN									
Query Match 57.4%; Score 211.4; DB 5; Length 245;									
Best Local Similarity 95.5%; Pred. No. 1.1e-57;									
Matches 237; Conservative 0; Mismatches 9; Indels 2; Gaps 2;									
Qy	76	atcttctctcaagtccctggccctgttgaatctatctgtgataacactggagctgaactcc	135						
Db	1	ATCTTCTCTCAGTGGCCCTGGCCNCTGTGAGTCTATCTGTTAACTGAGCACTGGAGCTGACTCC	60						
Qy	136	ctgggaagagagcccaaatgttacaatgaacttaataatgaatgcaccaagatatagacct	195						
Db	61	CTGGGAAGAGAGGCCAAATGTTACAATGAACITTAATGATGCACCAAGATAATGACCTT	120						
Qy	196	gtctgtgggactgaatgaataacttaccatgaatgcgaatgcgtgtatg-ttttgaagctc	253						
Db	121	GTCGTGGGACTGATGAAATACCTTATCCCAATGAATGCCGTGTATGTTTITGAAATC	180						
Qy	254	ggaacgcagactcttaccctcattcacaataatcgagcccttgctgaagaccagagtttt	313						
Db	181	GGAAACGCCAGACTTCTATCCCTATTCATCAAAAATCTGGCCCTTNCIGAAACAGGGTIT	240						
Qy	314	gaataccc 321							
Db	241	NAAAAATCC 248							
RESULT 9									
E01725									
LOCUS	E01725 179 bp DNA linear PAT 29-SEP-1997								
DEFINITION	Synthetic DNA encoding human pancreatic secretory trypsin inhibitor (PSII).								
ACCESSION	E01725								
VERSION	E01725.1 GI:2169978								
KEYWORDS	JP 1988267289-A/1.								
SOURCE	synthetic construct.								
ORGANISM	artificial sequence.								
REFERENCE	1 (bases 1 to 179)								
AUTHORS	Obata,O., Shin,M., Kikuchi,N. and Teraoka,H.								
TITLE	NOVEL PRODUCTION OF PROTEIN								
JOURNAL	Patent: JP 1988267289-A 1 04-NOV-1988;								
COMMENT	SHIONOGI & CO LTD								
OS	Artificial gene								
OC	Artificial sequence: Genes.								
PN	JP 1988267289-A/1								
PD	04-NOV-1988								
PF	05-OCT-1987 JP 1987253437								
PR	14-OCT-1985 JP 86P 245048, 23-DEC-1985 JP 86P 314503 PI								
OBAPA	OSAMU, SHIN MASARU, KIKUCHI NORIHISA, TERAOKA HIROSHI FC								
C12P21/00	C12N15/00.(C12P21/00,C12R1:19):								
CC	strandedness: Double;								
CC	topology: Linear;								
CC	hypothetical: No;								
CC	anti-sense: No;								
FH	Key Location/Qualifiers								
FH	Key								
FT	CDS								
FT	mat_peptide 8..175 /product='synthetic human psi-1' FT								
FT	Location/Qualifiers								
FT	1..179								
FT	/organism='synthetic construct'								
FT	/db_xref='taxon:32630'								
BASE COUNT	54 a 37 c 41 g 47 t								
ORIGIN									
Query Match 45.9%; Score 168.8; DB 6; Length 179;									
Best Local Similarity 98.8%; Pred. No. 7.9e-44;									
Matches 170; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
Qy	130	gactccctgggaagagcccaaatgttacaatgaacttaataatgaatgcgaatgcacaaagatatat	189						
Db	9	GACTCCCTGGGAAGAGAGCCCAAATGTTACAATGAACITTAATGGAATGCCCAAGAATAAT	57						
Qy	190	gacctctctgtgggactgaatgaataacttaccatgaatgcgaatgcgtgtatgtttttaa	249						
Db	62	GACCCCTGTCGTGGGACTGATGGAATACTTATCCCAATGAATGCCGTGTATGTTTIGAA	127						
Qy	250	agtcggaagagccagacttctaccctcattcacaataatcgagcccttgctgaag	301						
Db	128	AATCGGAAAGCCGACACTTCTATCCTCATCAAAAATCTGGCCCTTGCTGAG	179						
RESULT 10									
E02456									
LOCUS	E02456 179 bp DNA linear PAT 29-SEP-1997								
DEFINITION	DNA encoding human modified pancreatic secretory trypsin inhibitor (human PSII).								
ACCESSION	E02456								
VERSION	E02456.1 GI:2170687								
KEYWORDS	JP 1990150282-A/2.								
SOURCE	synthetic construct.								
ORGANISM	artificial sequence.								
REFERENCE	1 (bases 1 to 179)								
AUTHORS	Yoshida,N., Kikuchi,N., Shin,M. and Teraoka,H.								
TITLE	MODIFIED HUMAN PSII								
JOURNAL	Patent: JP 1990150282-A 2 08-JUN-1990;								
COMMENT	SHIONOGI & CO LTD								
OS	Artificial gene								
OC	Artificial sequence: Genes.								
PN	JP 1990150282-A/2								
PD	08-JUN-1990								
PF	11-OCT-1988 JP 1988255580								
PR	19-JUL-1988 JP 88P 181316								
PI	YOSHIDA NOBUO, KIKUCHI NORIHISA, SHIN MASARU, TERAOKA HIROSHI								
PC	C12N15/12.C07K7/10//C07H21/04,C12P21/02,(C12P21/02,C12P1:19):								
PC	C07K99:00;								
CC	strandedness: Double;								
CC	topology: Linear;								
CC	hypothetical: No;								
CC	anti-sense: No;								
CC	*source: clone-pUC13-PSII;								
FH	Key Location/Qualifiers								
FH	Key								
FT	5'UTR								
FT	CDS								
FT	mat_peptide 5..175 /product='Met-wild type human PSII' FT								
FT	Location/Qualifiers								
FT	179..183								
FT	mutation								
FT	replace(132,'a')								
FT	/note='Gln(42)-PSII'								
FT	mutation								
FT	replace(135,'c')								
FT	/note='Thr(43)-PSII'								
FT	mutation								
FT	replace(137,'a')								
FT	/note='Ser(44)-PSII'								
FT	misc_feature 8..175								
FT	/product='wild type human PSII'.								
FT	Location/Qualifiers								
FT	1..179								
FT	/organism='synthetic construct'								
FT	/db_xref='taxon:32630'								
BASE COUNT	54 a 37 c 41 g 47 t								
ORIGIN									

BASE COUNT	103 a	107 c	88 g	111 t	11 others
ORIGIN	/db_xref="taxon:9506"				
Query Match	44.0%	Score 162.2	DB 6	Length 420	
Best Local Similarity	95.2%	Pred. No. 1.2e-41			
Matches 175	Conservative	0	Mismatches	6	Indels 1
Gaps	1				
Qy	150	aatgaacttaagtatgacccaagatataatgacccctgtctgtggaactgaatgaataac	218		
Db	200	AANGAATTAATGATGACCAACCAAGATATATGACCCNGCTGTGGGACTGATGGAAATAC	141		
Qy	219	ttatcccaatgaatgctgttatgtttttaaagtcgaaacgcacgaactcttatctctcat	278		
Db	140	TTATCCCAATGAATGCTGTATGTGTTTGAATAATCGGAAGCCAGACATTCATCNICAT	81		
Qy	279	tcaaaatctggccctgtctgagaaacaaagtttttaaatcccatcaagtcaccuccaag	338		
Db	80	TCAAAAATCTGGGCCCTTCCTGAGAACCAAGGTTTGAATCCCATCAGGTCACCCGAAAG	21		
Qy	339	cc 340			
Db	20	AC 19			
RESULT 14					
AP059723					
LOCUS	AP059723	Sequence 26	from patent US 5840518	linear	FBI 29-SEP-1999
DEFINITION	Sequence 26 from patent US 5840518				
ACCESSION	AP059723				
VERSION	AP059723.1 GI:5986173				
KEYWORDS	Unknown				
SOURCE	Unknown				
ORGANISM	Unclassified				
REFERENCE	1 (bases 1 to 302)				
AUTHORS	Morishita,H., Kanamori,I. and Nobuhara,M.				
TITLE	DNA fragment, vector containing the DNA fragment, transformant transformed with the vector and process for producing protein using the vector				
JOURNAL	Patent: US 5840518-A 26 24-NOV-1998				
FEATURES	source				
BASE COUNT	97 a	64 c	58 g	73 t	
ORIGIN	1..302				
Query Match	44.0%	Score 162	DB 6	Length 302	
Best Local Similarity	91.9%	Pred. No. 1.3e-41			
Matches 171	Conservative	0	Mismatches	15	Indels 0
Gaps	0				
Qy	120	cactggagctgactccctctggaagagagagcccaaatgtttacaaatgaacttaagtatgac	179		
Db	116	CTCGGAATGACCTCCCTAGGTCCGGAGGCCAAATGTTACAAATGAATTAATGGAATGAC	175		
Qy	180	caagatatgaactctctgtggaactgaatgaataactttatcccaatgaatgctgtt	239		
Db	176	CAAGATAATGACCTCTGTGTGGGACTGATGGAATACTTATCCCAATGAATGCGTGT	235		
Qy	240	atgttttgaagtcgaaacgcagactctctctcattccaaatctgtgcccctgtg	299		
Db	236	ATGTTTGAATAATCGGAAGCCGACATCGATCCTCATTCACAAATCTGGGCTTGCTG	295		
Qy	300	agaacc 305			
Db	296	AGGATC 301			
RESULT 15					
E08411					
LOCUS	E08411	Sequence 302	bp	DNA	linear FBI 29-SEP-1997

Search completed: July 29, 2002, 17:55:02
Job time: 5168 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Fun on:

July 29, 2002, 16:28:54 : Search time 1928.76 seconds
(without alignments)
5055.977 Million cell updates/sec

Title: US-09-880-107-2492

Perfect score: 466

Sequence:

Scoring table: 1 atcaagaacacatagattc.....gggtaaggtggga:ggctctt 466

IDENTITY_MUC

Gapop 10.0, Gapext 1.0

Searched:

1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3555312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb.ba:*

2: gb.htg:*

3: gb.in:*

4: gb.om:*

5: gb.ov:*

6: gb.pat:*

7: gb.ph:*

8: gb.pl:*

9: gb.pr:*

10: gb.ro:*

11: gb.sts:*

12: gb.sy:*

13: gb.un:*

14: gb.vi:*

15: em.ba:*

16: em.fun:*

17: em.hum:*

18: em.in:*

19: em.mu:*

20: em.or:*

21: em.ov:*

22: em.pat:*

23: em.ph:*

24: em.pl:*

25: em.ro:*

26: em.sts:*

27: em.un:*

28: em.vi:*

29: em.htg:*

30: em.htg.in:*

31: em.htg.in:*

32: em.htg.other:*

33: em.htg.in:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	466	100.0	456	5	AX329950	Sequence
2	400.2	85.5	155874	2	AL645936	Homo sapi
3	398.6	85.5	971	9	AF123050	Homo sapi
4	398.6	85.5	134292	9	HS271M21	Human DNA
5	398.6	85.5	211591	2	AL662826	Homo sapi
6	397	85.2	777	9	HS21UB1QU	Homo sapi
7	397	85.2	41069	2	AC004179	Homo sapi
8	397	85.2	129567	2	AL671618	Homo sapi
9	397	85.2	129806	9	AC006137	Homo sapi
10	395.4	84.8	791	9	BC012472	Homo sapi
11	394.4	71.8	474	6	AX237229	Homo sapi
12	394.4	71.8	474	6	AX237229	Homo sapi
13	323.4	69.4	453	6	AX237670	Sequence
14	323.4	69.4	453	6	AX237670	Sequence
15	323.4	69.4	453	6	AX237670	Sequence
16	323.4	69.4	453	6	AX237670	Sequence
17	323.4	69.4	453	6	AX237670	Sequence
18	323.4	69.4	453	6	AX237670	Sequence
19	323.4	69.4	453	6	AX237670	Sequence
20	323.4	69.4	453	6	AX237670	Sequence
21	323.4	69.4	453	6	AX237670	Sequence
22	323.4	69.4	453	6	AX237670	Sequence
23	323.4	69.4	453	6	AX237670	Sequence
24	323.4	69.4	453	6	AX237670	Sequence
25	323.4	69.4	453	6	AX237670	Sequence
26	323.4	69.4	453	6	AX237670	Sequence
27	323.4	69.4	453	6	AX237670	Sequence
28	323.4	69.4	453	6	AX237670	Sequence
29	323.4	69.4	453	6	AX237670	Sequence
30	323.4	69.4	453	6	AX237670	Sequence
31	323.4	69.4	453	6	AX237670	Sequence
32	323.4	69.4	453	6	AX237670	Sequence
33	323.4	69.4	453	6	AX237670	Sequence

ALIGNMENTS

RESULT 1
AX329950
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 459 13-DEC-2001;
Avalon Pharmaceuticals (US)
Location/Qualifiers
1: 466
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COVI 119 a 130 c 89 g 129 t

AX329950 Sequence 456 bp DNA linear PAT 09-JAN-2002

Sequence 459 from Patent WO0194629.

AX329950

AX329950.1 GI:18102928

human

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (sites)

Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,

Horrikan, S., Soppet, D.R. and Weaver, Z.

Cancer gene determination and therapeutic screening using signature

gene sets

Patent: WO 0194629-A 459 13-DEC-2001;

Avalon Pharmaceuticals (US)

Location/Qualifiers

1: 466

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COVI 119 a 130 c 89 g 129 t

ORIGIN

Query Match 100.0%; Score 466; DB 5; Length 466
 Best Local Similarity 100.0%; Pred. No. 1.6e-131; Indels 0; Gaps 0;
 Matches 466; Conservative 0; Mismatches 0

Qy 1 atcaagaacacatagagtcgggcaataatactctcctaccatccacccaacttta 60
 |||||
 Db 1 ATCAAGAAACATAGAGTTCGGCAATATACCTTCATCTACCCATCCCAACCAATCTTA 60
 |||||
 Qy 61 ctctactcattctctcatttaatttgggaataatcagaagaatatttcctgaagta 120
 |||||
 Db 61 CTCTACTCATCTCATTTCTCAITTAATTTTGGGAAATCATCAGAAGATGTTTCCTT 120
 |||||
 Qy 121 agagattaaagaataaagctttttgacccctgcacaccccccattccacaggtttcac 180
 |||||
 Db 121 AGAGATTAAAGAATAAAGCTTTTTCACCCCTGCGCAACACCCCATGCCAGGCTT 180
 |||||
 Qy 181 cctccatacaataacatgccagaagagtaagttaccctttctgaatgccgaattccc 240
 |||||
 Db 181 CCTCCAAATACAATAACATGCGAGGAAGTAAGTTGCCCTTTCTGAIGCGGTAAT 240
 |||||
 Qy 241 atcatcttcccatcttccatctcctttccattgcgaagtcacaattcgggtctctggat 300
 |||||
 Db 241 ATCATCTTCCCATCTTCCAGTCTCTTTCATTTGCAAGTCACAACTGGGCTCTCA 300
 |||||
 Qy 301 taccctgtcttaagctcgaatcattgtcttccattgcacacacacacacacacacacac 360
 |||||
 Db 301 TATACCCGCTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 360
 |||||
 Qy 361 ctggagagagtgctcttgcctcattcactcactcactcactcactcactcactcactcact 420
 |||||
 Db 361 CTGGAGAGGTGCTCTTTCCTCATCTCCTCATCTCCTCATCTCCTCATCTCCTCAT 420
 |||||
 Qy 421 catcactggggtctcaccacttccaggggtaagtgagtgagtttt 466
 |||||
 Db 421 CATCACTGGGCTTCACTCACTTTCAGGGTAAGGTGGATGGTCTT 466
 |||||

RESULT 2

AL645936 155874 bp DNA linear 813 31-JAN-2002
 LOCUS Homo sapiens chromosome 6 clone XHbac-126D10, *** SEQUENCING IN
 PROGRESS ***, in ordered pieces.
 AL645936
 HG: HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLIOP.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (sites)
 Leongamornlert.D.
 Direct Submission
 Submitted (25-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
 On Feb 1, 2002 this sequence version replaced gi:1615238.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquerry@sanger.ac.uk
 ----- Project Information
 Center project name: BPG126D10
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; 108752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Consensus quality: 155824 bases at least Q40
 Consensus quality: 155864 bases at least Q30
 Consensus quality: 155872 bases at least Q20

COMMENT

Insert size: 155874; sum-of-contigs
 Insert size: 162470; 3.9% error; agarose-fp
 Quality coverage: 10.66x in Q20 bases; sum-of-contigs Quality
 coverage: 10.30x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 Location/Qualifiers
 1. 155874
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="XHbac:126D10"
 /clone_11b="CHORI-501"
 1. 155874
 /note="assembly_fragment:03519"
 BASE COUNT 43601 a 36123 c 34182 g 41968 t
 ORIGIN
 Query Match 85.9%; Score 400.2; DB 2; Length 155874;
 Best Local Similarity 98.5%; Pred. No. 3e-111;
 Matches 446; Conservative 0; Mismatches 3; Indels 4; Gaps 4;
 Qy 1 atcaagaacacatagagtcgggcaataatactctcctaccatccacccaacttta 60
 |||||
 Db 6691 ATCAAGAAACATAGAGTTCGGCAATATACCTTCATCTACCCATCCCAACCAATCTTA 6749
 |||||
 Qy 61 ctctactcattctctcatttaatttgggaataatcagaagaatatttcctgaagta 120
 |||||
 Db 6750 CTCTACTCATCTCATTTCTCAITTAATTTTGGGAAATCATCAGAAGATGTTTCCTT 120
 |||||
 Qy 121 agagattaaagaataaagctttttgacccctgcacaccccccattccacaggtttcac 180
 |||||
 Db 6810 AGAGATTAAAGAATAAAGCTTTTTCACCCCTGCGCAACACCCCATGCCAGGCTT 180
 |||||
 Qy 181 cctccatacaataacatgccagaagagtaagttaccctttctgaatgccgaattccc 240
 |||||
 Db 6870 CTCTCAATACAATAACATGCGAGGAAGTAAGTTGCCCTTTCTGAIGCGGTAAT 240
 |||||
 Qy 241 atcatcttcccatcttccatctcctttccattgcgaagtcacaattcgggtctctggat 300
 |||||
 Db 6930 ATCACTTCCCATCTTCCAGTCTCTTTCATTTGCAAGTCACAACTGGGCTCTCA 300
 |||||
 Qy 361 taccctgtcttaagctcgaatcattgtcttccattgcacacacacacacacacacacac 360
 |||||
 Db 6989 TATACCCGCTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 360
 |||||
 Qy 361 ctggagagagtgctcttgcctcattcactcactcactcactcactcactcactcactcact 420
 |||||
 Db 7049 CT-GGAGGAGGTGCTCTTTCCTCATCTCCTCATCTCCTCATCTCCTCATCTCCTCAT 420
 |||||
 Qy 421 catcactggggtctcaccacttccaggggtaagtgagtgagtttt 453
 |||||
 Db 7108 CATCACT-GGGCTTCACTCACTTTCAGGGTAAGG 7139
 |||||
 RESULT 3
 AF123050/2
 LOCUS Homo sapiens ubiquitin-like protein FA110 (FA110) mRNA, complete
 DEFINITION cds.
 ACCESSION AF123050
 VERSION AF123050.1 GI:5821714
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 971)
 REFERENCE
 Liu,Y.C., Pan,J., Zhang,C., Fan,W., Collinge,M., Bender,J.R. and


```
repeat_region 5815..5853
/notes="MIR repeat: matches 108..146 of consensus"
repeat_region 7324..7786
/notes="HCRV15 repeat: matches 596..1091 of consensus"
repeat_region 7857..8024
/notes="LIM4 repeat: matches 3415..3586 of consensus"
repeat_region 8151..8565
/notes="IHE1C repeat: matches 1..371 of consensus"
repeat_region 8504..8910
/notes="Alusx repeat: matches 2..306 of consensus"
repeat_region 9002..9354
/notes="MLT1B repeat: matches 1..390 of consensus"
repeat_region 9717..9795
/notes="MIR repeat: matches 20..98 of consensus"
repeat_region 10027..10047
/notes="LIMB6 repeat: matches 5767..5786 of consensus"
repeat_region 10048..10493
/notes="MLT1C repeat: matches 1..465 of consensus"
repeat_region 10494..10848
/notes="LIMB5 repeat: matches 5786..6145 of consensus"
repeat_region 11216..11356
/notes="L2 repeat: matches 2554..2703 of consensus"
repeat_region 11386..11537
/notes="MIR repeat: matches 7..167 of consensus"
repeat_region 11610..11785
/notes="LIMC2 repeat: matches 4759..4931 of consensus"
repeat_region 11877..12095
/notes="Alusg1 repeat: matches 1..308 of consensus"
repeat_region 12036..13044
/notes="LIMC2 repeat: matches 4931..5829 of consensus"
repeat_region 13093..13476
/notes="LIMB1 repeat: matches 5728..6118 of consensus"
repeat_region 13916..13921
/notes="MIR repeat: matches 69..147 of consensus"
repeat_region 14700..14834
/notes="MIR repeat: matches 89..226 of consensus"
repeat_region 16392..16688
/notes="MER9 repeat: matches 215..511 of consensus"
repeat_region 16994..17015
/notes="L1 copies 2 mer aa 100% conserved"
misc_feature 17498..17507
/notes="match: GSS: Em:A0883063"
misc_feature 17524..17673
/notes="match: GSS: Em:B72414"
repeat_region 17750..18033
/notes="LIMC/D repeat: matches 5479..5757 of consensus"
misc_feature 18059..18568
/notes="match: GSS: Em:A0743745"
repeat_region 18234..19882
/notes="MER4-internal repeat: matches 2701..4900 of consensus"
repeat_region 19883..20181
/notes="AluJb repeat: matches 1..292 of consensus"
repeat_region 20182..20687
/notes="MER4-internal repeat: matches 2198..4701 of consensus"
repeat_region 20675..20796
/notes="MER4-internal repeat: matches 2341..4454 of consensus"
repeat_region 20777..20853
/notes="MER4-internal repeat: matches 1034..1114 of consensus"
repeat_region 20854..21376
/notes="LIR35 repeat: matches 1..612 of consensus"
repeat_region 21377..22220
/notes="MER4-internal repeat: matches 121..1037 of consensus"
repeat_region 22222..22934
/notes="MER50 repeat: matches 1..711 of consensus"
repeat_region 22959..23269
/notes="AluJb repeat: matches 1..307 of consensus"
repeat_region 23272..23353
/notes="MER4-internal repeat: matches 1..84 of consensus"
```

```
repeat_region 23357..23776
/notes="MER4C repeat: matches 1..465 of consensus"
misc_feature 25589..27225
/notes="match: GSS: Em:AQ195255
match: STS: Em:G61733"
repeat_region 28927..29156
/notes="SN5 repeat: matches 1..235 of consensus"
repeat_region 30633..30690
/notes="29 copies 2 mer gt 74% conserved"
repeat_region 31019..31335
/notes="AluJb repeat: matches 1..311 of consensus"
repeat_region 31856..32451
/notes="LIM4 repeat: matches 4650..5281 of consensus"
repeat_region 32517..32893
/notes="LIM4 repeat: matches 3785..4144 of consensus"
repeat_region 32899..33125
/notes="LIMC repeat: matches 1787..2016 of consensus"
misc_feature 33093..33554
/notes="match: GSS: Em:AQ620121"
repeat_region 33207..33337
/notes="LIM4 repeat: matches 4708..4838 of consensus"
repeat_region 33458..33574
/notes="MIR repeat: matches 144..252 of consensus"
repeat_region 33679..34067
/notes="MSIA repeat: matches 1..426 of consensus"
gene complement(34169..35034)
/genes="dJ271M21.8"
CDS complement(34169..35034)
/genes="dJ271M21.8"
/notes="dJ271M21.8 (MAS1 oncogene pseudogene)
match: proteins: Sw:P30554 Sw:P2526 Sw:P04201 Sw:P35410"
/notes="codon_start=1
/pseudo
/evidence=not_experimental
35631..35754
/notes="L2 repeat: matches 2367..2511 of consensus"
repeat_region 35169..36279
/notes="MER81 repeat: matches 3..113 of consensus"
repeat_region 35455..36829
/notes="MLT1B repeat: matches 1..374 of consensus"
repeat_region 37958..38292
/notes="LIMB5 repeat: matches 5817..6165 of consensus"
repeat_region 39291..39590
/notes="AluSp repeat: matches 1..301 of consensus"
repeat_region 40192..41536
/notes="LIM4 repeat: matches 2220..3639 of consensus"
repeat_region 41538..41700
/notes="LIM4C repeat: matches 1140..1302 of consensus"
repeat_region 41701..42011
/notes="AluYa5 repeat: matches 1..311 of consensus"
repeat_region 42012..43297
/notes="LIM4C repeat: matches 1301..2216 of consensus"
repeat_region 43281..43540
/notes="LIM45 repeat: matches 5814..6059 of consensus"
repeat_region 43538..43961
```

```
Query Match 85.5% Score 398.6 DB 9 Length 134292
Best Local Similarity 98.2% Pred. No. 9.2e-111
Matches 445 Conservative 0 Mismatches 4 Indels 4 Gaps 4
```

```
Qy 1 tcaagaacatagagtcgggcaataattcttctaccatccaccacccaatttta 60
|||||
Db 52163 ATCAAGAAACATAGAGTT-GGGCAATAIACITCATCTACCCATCCACCAATCITA 52161
Qy 61 tctactcatctctctcattataatttgggaatacatcaagaatgtgttcgttgagta 120
|||||
Db 52162 CTCIACATCACTCAITCTCAITTAATTTGGGAATCATCAGAGATGTGTTCGTTAGTA 52221
Qy 121 agagataaaagaaataagctttttgacccctgccacaacccctgccagggtggtcac 180
|||||
Db 52222 AGAGATTAAAGAAATAAGCTTTTIGACCCCTGCCAACACCCCATGCCAGGGTGGTGCAC 52281
Qy 191 actcaataacaataacatgccagagagagtaagtgtgcctttctgatccgtaatctgcc 240
```



```

REFERENCE          2 (bases 1 to 777)
AUTHORS            Bates, E.F.M.
TITLE              Direct Submission
JOURNAL            Submitted (17-APR-1997) E.F.M. Bates, Schering-Plough, Laboratory
                  for Immunological Research, 27 Chemin des Peupliers, BP11, 95571
                  Bardilly Cedex, FRANCE
FEATURES
  Source           1..777
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /cell_type="dendritic cells"
                  /codon_start=1
                  /product="diubiquitin"
                  /protein_id="CAA73200.1"
                  /db_xref="GI:2546964"
                  /db_xref="SPTREMBL:O15205"
                  /translation="MAPNASCILCVHVRSEENDLMTFDANPYDSVKIKHEVSKIKVP
                  VQDVLLGSKILKPRSLSSGIDKEKIIHLTKVKFSDEELPLFLVESDEAKRH
                  LQVRSSVAOVKAMIEIKTGIPIQIVICNGKRLDEKRWADYGRKGNLFLAS
                  YCIGG"
  polyA_signal     223 a 153 c 195 g 206 t
                  748..753
  CDS
  polyA_signal     223 a 153 c 195 g 206 t
                  748..753
  ORIGIN
Query Match      85.2%; Score 397; DB 9; Length 777;
Best Local Similarity 98.0%; Pred. No. 2.7e-110;
Matches 444; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

OY 1 atcaagaacatagattcgggaataatacttcattcatccaccaccaccccaaatctta 60
DB 591 ATCAAGAAACATAGATT-GGCAATATATCTTCTACCCATCCACCCCAATCTTA 633
OY 61 ctctactcatctctcattcattatatttgggaataatcatcagaagatgtttcgttgagta 120
DB 632 CTCTACTCATCTCATCTCATTAATTTTGGGAATCATCAGAAGATGTTTCGTIGAGTA 573
OY 121 agagattaaagaataagctttttgacctgtccacacaccccatgccaggtggtcac 180
DB 572 AGAGATTAAAGAAATAAGCTTTTIGACCCCTGCCACACACCCCATCCCGSGTGGTCAC 513
OY 181 ctctcaatacaataacatgccaggaagaagttaagttgccctttctgattgccataatgcc 240
DB 512 CTCTCAATACATATAGATGCCAGGAAGATTAAGTTGCCCTTTCATGCGGTAATCTGCC 453
OY 241 atcatcttccatcttccagtcctccttccattgcaagtcacaactcgggtctcaggat 300
DB 452 ATCATCTTCCCATCTTCCAGTCT-CTTTCCATTCGAAGTCACAATCTGGGTCTCAGGGAT 334
OY 301 tataccgctttagctcgtcattgttttcaactgtgcccactgaagctgggaccttcgcac 360
DB 393 TATACCGCTTAGTCTCGATCATTTGCTTTCATTTGSCCATGAGCTGGACCTTCGCAC 334
OY 361 ctggagagaggtgcctctttaccctcatcactgactccacagaagaacaaagtcagctcct 420
DB 333 CT-GGAGGAGGTGCCTCTTTGGCTTCATCACCCTGACCTGACACAGAAACAAAGSGAGCTCT 275
OY 421 catcactggggttccaccacttccagggttaag 453
DB 274 CATCACT-GGGCTTCACCACATTTTCAGGGTAAGG 243

RESULT 7
AC004179
LOCUS              Homo sapiens chromosome 6 clone UWGC:CB map p21. *** SEQUENCING IN
DEFINITION         AC004179 41069 bp DNA linear HTG 23-FEB-1998
PROGRESS ***; 1 ordered pieces.
ACCESSION          AC004179
VERSION            AC004179.1 GI:2905881
KEYWORDS           HTG; HTGS_PHASE2.
SOURCE             human.
ORGANISM           Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 41069)

Janer, M., Guillaudoux, I., Vu, O., Kutyavin, T., Harter, H. and Geraghty, D.E.

Large scale sequence analysis of the human MHC class I region

Unpublished (1998)

Fred Hutchinson Cancer Research Center

The Clinical Research Division

1100 Fairview Ave. N., P.O. Box 19024

Seattle, WA 98109-1024

2 (bases 1 to 41069)

Geraghty, D.E. and Olson, M.V.

Direct Submission

Submitted (23-FEB-1998) Human Genome Center, University of

Washington, Box 352145, Seattle, WA 98195, USA

University of Washington Human Genome Center

Box 352145 Seattle, WA 98195

Contact: Daniel E. Geraghty (geraghty@fhcrc.org).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 41069: contig of 41069 bp in length.

Location/Qualifiers

1..41069

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="6"

/map="p21"

/clone="UWGC:CB"

BASE COUNT 12023 a 9129 c 8744 g 11173 t

ORIGIN

Query Match 85.2%; Score 397; DB 2; Length 41069;

Best Local Similarity 98.0%; Pred. No. 2.7e-110;

Matches 444; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

OY 1 atcaagaacatagattcgggaataatacttcattcatccaccaccaccccaaatctta 60

DB 18419 ATCAAGAAACATAGATT-GGCAATATATCTTCTACCCATCCACCCCAATCTTA 18477

OY 51 ctctactcatctctcattcattatatttgggaataatcagaagatgtttcgttgagta 120

DB 18478 CTCTACTCATCTCATCTCATTAATTTTGGGAATCATCAGAAGATGTTGTTGAGTA 18537

OY 121 agagattaaagaataaagctttttgacctgtccacaccccatgccaggtggtcac 180

DB 18538 AGAGATTAAAGAAATAAGCTTTTIGACCCCTGCCAACACCCCATCCCGSGTGGTCAC 18597

OY 181 ctctcaatacaataacatgccaggaagaagttaagttgccctttctgattgccataatgcc 240

DB 18598 CTCTCAATACAAATAAGATGCCAGGAGAGTAAGTTGCCCTTTCTGATGCCGTAAATCTGCC 18657

OY 241 atcatcttcccatcttccagttcctctcttccattgcaagtccacaaatcgtggtctcaggat 300

DB 18659 ATCACTTCCCATCTTCCAGTCT-CTTTCCATTGCAAGTCACAATCTGGTCTCAGGGAT 18716

OY 301 tataccgctttagctcgtatcattgtcttccattgtgacctgaagcttgaccttcgcac 360

DB 18717 TATACCGCTTCTAGTCTCGATCATTTGCTTTCACCTTGTGCCACTGAGCTGGACCTTCGCAC 18776

OY 361 ctggagagaggtgcctctttgacctcatcactcactccacagaagaacaaagtcagctcct 420

DB 18777 CT-GGAGGAGGTGCCTCTTTGCCCTTCATCCTGACTGACTCCACAGAAACAAAGSGAGCTCT 18835

```

O> 421 caccactgggcttcaccactttcagaggtaag 453
||||| ||||||| ||||||| ||||||| |||
Db 18836 CATCACT-GGGCTTACACACATTCAGGGTAAGG 18857

RESULT      8
AL571618/c  129567 bp  DNA  linear  HIG 01-FEB-2002
LOCUS      Homo sapiens chromosome 6 clone XXbac-140P7, *** SEQUENCING IN
DEFINITION  PROGRESS ***, 2 unordered pieces.
ACCESSION  AL571618
VERSION    AL571618.2 GI:18491494
KEYWORDS  HIG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HIGS_FULLTOP.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (sites)
            Direct Submission
            Griffiths.C.
            Submitted (31-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquer@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            On Feb 4, 2002 this sequence version replaced gi:18477174.
            ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquer@sanger.ac.uk
            ----- Project Information
            Center project name: bpg140P7
            ----- Summary Statistics
            Sequencing program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 129267 bases at least Q40
            Consensus quality: 129282 bases at least Q30
            Consensus quality: 129363 bases at least Q20
            Insert size: 129467; sum-of-contigs
            Insert size: 135304; 5.8% error; agarose-fp
            Quality coverage: 10.92x in Q20 bases; sum-of-contigs Quality
            coverage: 10.45x in Q20 bases; agarose-fp
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            * 1 103125: contig of 103125 bp in length
            * 103126 103225: gap of 100 bp
            * 103226 129567: contig of 26342 bp in length.
            Location/Qualifiers
              1..129567
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="6"
                /clone="XXbac-140P7"
                /clone_lib="CHORI-501"
                /clone_1="1"
                /note="assembly fragment:02311"
                fragment_chain:1
                  clone_end:17
                  vector_side:left"
            misc_feature
              1..103125
                /note="assembly fragment:00372"
                fragment_chain:1
                  clone_end:SP6
                  vector_side:right"

BASE COUNT  36114 a 27532 c 28517 g 37304 t 100 others

```

ORIGIN

```

Query: Watch      85.2%   Score 397;   DB 2;   Length 129567;
Best Local Similarity 98.0%;   Pred. No. 2.8e-110;
Matches 444;   Conservative 0;   Mismatches 5;   Indels 4;   Gaps 4:

Q> 1 atcaagaacacatagattcgggcaataatacttctaccctaccctcccaaatctta 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4382 ATCAAGAAAATAGAGTT-GGGCAATATATCTTCACTCTACCCTACCCCAAATCTTA 4324
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Q> 61 ctctactcatctctctctctctctctctctctctctctctctctctctctctct 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4323 CTCTACTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4264
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Q> 121 agagattaaaagaataagctttttgacccctcgcacacccctcgcacacccctcgcac 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4263 AGAGATTAAAAGAAATAAGCTTTTIGACCCCTGCCAACCCCTACCCAGGGTGGTCAC 4204
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Q> 181 cctccatacaataacatgcacagagaagtaagtgccttctctctctctctctctctct 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4203 CTTCCAAATACAAATAAGATGCCAGGAGTAAGTTGCGCTTTCTGATGCCGTAACTGCC 4144
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Q> 241 atcatctcccatctctccagctctctctctctctctctctctctctctctctctct 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4143 AICATCTTCCATCTTCCAGTCT-CTTTCCATTGCAAGTCACAAATCTGGGTCTCAGGGAI 4085
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Q> 301 tataccgcttagctctcgatctgtctctctctctctctctctctctctctctctctct 360
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4084 TATACCCGCTTAGTCTCGATCATGTCTTTCACCTGACCTGACCTGACCTGACCTGAC 4025
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Q> 361 ctggagagagtgctctctctctctctctctctctctctctctctctctctctctct 420
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4024 CT-GGAGGAGGIGGCTCTTTTGGCTTCATCACCCTGACTCCACAAAGAACAGGGAGCTCT 3956
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Q> 421 catcctgggcttcacacactttcagaggtaag 453
||||| ||||||| ||||||| ||||||| |||
Db 3965 CATCACT-GGGCTTACACACATTCAGGGTAAGG 3934

RESULT      9
AC006137/c  129806 bp  DNA  linear  PRI 19-OCT-2001
LOCUS      Homo sapiens clone SCB-254N2 (UMGC:rq254N02) from 6p21, complete
DEFINITION  sequence.
ACCESSION  AC006137
VERSION    AC006137.3 GI:16271961
KEYWORDS  HIG.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 129806)
            Janer,M., Guillaudoux,T., Vu,Q., Kutayavin,T., Harter,H. and
            Geraghty,D.E.
            Large scale sequence analysis of the human MHC class I region
            Unpublished
            Fred Hutchinson Cancer Research Center
            The Clinical Research Division
            1100 Fairview Ave. N., P.O. Box 19024
            Seattle, WA 98109-1024
            2 (bases 1 to 129806)
            Geraghty,D.E. and Olson,M.V.
            Direct Submission
            Submitted (05-DEC-1998) Human Genome Center, University of
            Washington, Box 352145, Seattle, WA 98195, USA
            University of Washington Human Genome Center
            Box 352145, Seattle, WA 98195
            Contact: Daniel E. Geraghty (geraghty@fhcrc.org)
            3 (bases 1 to 129806)
            Kaul,R.K. and Haugen,E.D.
            Direct Submission
            Submitted (24-AUG-2000) Genome Center, University of Washington.
            JOURNAL

```

REMARK Box 352145, Seattle, WA 98195, USA
 University of Washington Genome Center
 Box 352145 Seattle, WA 98195
 Contact: Pajinder Kaul (rkkaul@u.washington.edu)
 4 (bases 1 to 123806)
 Kaul, P.K. and Haugen, E.D.
 Direct Submission
 Submitted (12-OCT-2001) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 On Oct 19, 2001 this sequence version replaced gi:9887710.

Sequence Validation:
 This sequence has been validated by Multiple Complete Digest
 Mapping. Comparison of the experimentally derived map digest
 fragments with sequence-predicted fragments is given below.
 Small fragments below a variable cutoff (approximately 400-600bp)
 are not mapped and hence do not appear in the table. There are no
 significant remaining discrepancies between the experimental and
 predicted values. Uniquely ordered fragment groups are separated
 by dashed lines.

BglII		HindIII		NotI	
Map	Seq	Map	Seq	Map	Seq
739.54	740.00	2519.25	2440.00	5115.00	5097.00
3935.09	3908.00	4420.00	4411.00	6536.33	6521.00
5360.62	5319.00	5049.00	5001.00	5353.67	5202.00
4507.12	4455.00	1033.12	1012.00	10074.25	9987.00
1288.62	1303.00	2099.75	2064.00	1747.83	1751.00
12113.75	12356.00	17212.25	18083.00	10342.00	10215.00
2793.33	2676.00	10608.42	10467.00	4566.08	4599.00
4379.83	4328.00	1932.92	1925.00	2849.83	2850.00
1286.75	1280.00	12731.50	12500.00	1828.83	1831.00
879.17	863.00	1584.50	1580.00	3587.17	3722.00
1815.33	1826.00	819.00	788.00	4181.75	4110.00
11414.25	11761.00	2619.67	2569.00	10257.92	10055.00
6029.33	5933.00	2619.67	2581.00	5413.17	6349.00
6029.33	6005.00	953.00	954.00	1371.00	1370.00
5273.83	5308.00	1702.83	1697.00	1132.00	1121.00
15738.17	15951.00	1244.08	1238.00	5435.00	5587.00
5695.83	6654.00	3764.58	3791.00	3143.81	3146.00
2829.33	2792.00	2619.67	2605.00	7632.25	7732.00
731.67	723.09	1182.00	1181.00	4157.75	4197.00
1451.30	1461.00	4416.83	4379.00	754.07	745.00
3011.22	2991.00	12731.50	12909.00	2612.60	2609.00
8308.95	8430.00	4045.43	4133.00	6018.56	6104.00
5483.63	5517.00	1708.46	1697.00	5504.56	5411.00
1411.52	1408.00	2824.85	2828.00		
3036.94	3067.00	3283.08	3201.00		

FEATURES

Source
 Location/Qualifiers
 1..129806
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="5p21"
 /clone="UWGC:rg254N02"
 /clone_lib="Research Genetics"
 24582..24583
 /note="IS186 and target-site duplication removed from
 between bases 24682 and 26020"
 BASE COUNT 34075 a 28649 c 30144 g 36938 t
 ORIGIN

Query: Watch 85.2%; Score 397; DB 9; Length 129806;
 Best Local Similarity 98.0%; Pred No. 2.8e-110;
 Matches 444; Conservative 0; Mismatches 5; Indels 4; Gaps 4:
 QY 1 atcaagaacatagattgggcaatacttccatctcctaccatccccaccacaaatctta 60
 DB 101497 ATCAAGAACAATAGAGTT-GGGCAATATACITTCATCTACCCATCCCAACAAATCTTA 101429
 QY 11 ctctactcattcattctcatttaattttgggaatacatcagaagaatgtttcgttaagta 120
 DB 101429 CTCACCTACATCATTCATTAATTTTGGGAATCATCAGAAGATGTGTCGTTGAGTA 101369
 QY 121 agagattaaagaaataagctttttgacccctgccaacaccccatgccaggtgtgtcac 180
 DB 101262 AGAGATTAAGAAATAGCTTTTGGCCCTGCCAACACCCCAATCCCGAGGTGTGCAC 101309
 QY 161 ctccaataacaataacatcccaaggagagatgaagtgcctttctgatccgttaactgcc 240
 DB 101309 CTCCTCAATACAAATAGATGCCAGGAAGAGTAAGTTGCCCTTTCTGATGCCGTAATCTGCC 101249
 QY 241 atcaacttcccatcttccagatctcttccattgcaagtccaactctgggtctcaggat 300
 DB 101249 ATCACTTCCCATCTTCCAGTCT-CTTTCCATTGCAAGTCACAATCTGGGTCTCAGGAT 101190
 QY 361 tataccctcttagtctcgaatcattgcttccactgtgccactgagctgagaccttcgac 360
 DB 101169 TATAACGGCTTAGTCTCGATCATCTTCATCTTCTGCTGACCTGAGCTGAGCTTCGCAC 101130
 QY 361 ctggagagaggtgcctctttgctcatcactgcctgactcacaagaacaaagagacctct 420
 DB 101129 CT-GGAGGAGGTGCTCTTTGGCTCATCCTGACTCTCCACAAGAAACAGGCGACTCTCT 101071
 QY 421 catcactggggtctcaccactttcagggtgaag 453
 DB 101079 CATCACTTCCAGCTTTCAGGTTAAGG 101039

RESULT

LOCUS BC012472
 DEFINITION Hc-s sapiens, Similar to diubiquitin, clone MGC:21200
 IMAGE:4517233. mRNA, complete cds.
 ACCESSION BC012472
 VERSION BC012472.1 GI:15214679

KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 791)
JOURNAL Strausberg, P.
Direct Submission
Submitted (15-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cdapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalob@bcm.tmc.edu
Villalón, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 28 Row: 1 Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5821714.

FEATURES

Location/Qualifiers
1..791

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:21200 IMAGE:4517233"
/tissue_type="Bladder, transitional cell papilloma"
/clone_lib="NH_MGC_93"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT5"
25..522

CDS

/codon_start=1
/product="Similar to diubiquitin"
/protein_id="AAH12472.1"
/db_xref="GI:15214580"

/translation="MAPNASCCLCVHVRSEEDLMTFDAPHYDSYFKIKHVRSKTKVF
VQVLLGSKILKPRSSSYGIDKERTIHLTKVWPSEDELFELFVSGDEAKPH
LLQVRRSSVAQVKAMIEIKTGIIIPETQIVTCNGKRLDGKRMHADVIRKGNLLFLAC
YCIGG"

BASE COUNT 239 a 153 c 197 g 202 t

ORIGIN

Query Match 84.8%; Score 395.4; DB 9; Length 791;
Best Local Similarity 97.8%; Pred. No. 6.5e-110;
Matches 443; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
C 1 atcaagaacacatagagttcggaacataatacttctcaccatcccccaaaactta 60
D 697 ATCAAGAAACATAGAGTT-GGGCAATATATCTTACCTTACCCATCCAGCCAACTTA 639
C 61 ctctactctctctctcattcattattttgggaataatcatcagaagatgtgttcagta 120
D 638 CTCTACTCATCTCATCTCATTATTTTGGGAATATCATCAGAGATGTGTCTGATGTA 579
C 121 agagattaaagaataagctttttgacccctcccaacaccccatccccagggtgacac 180
D 578 AGAGATTAAAGAAATTAAGCTTTTTCACCCCTGCCAACACCCCATGCCAGGGTGTCAC 519
C 181 ctcccaatacaataacatccacagagatgaatttcccttttctcgaatgccgaatctgcc 240
D 518 CTCCCAATCAATAACATGCCAGAGAGTAAGTTGCCCTTTCTGATGCCGTAATTCGCC 459

C 241 atcatcttcccatcttccagttctcttccattgcaagtccacatctgggtctcaggat 300
D 458 ATCACTTCCCATCTTCAGTCT-CTTCCATTCAGTTCACATCTGGTCTCAGGAT 400
C 301 tataccctcttagctctcgatcatgttttcaacttggccactgagctgaccttcgac 360
D 399 TATACCCGTCTTAGTCTCGATCATGTCTTCACTTGTGCCACTGAGCTGGACCTTCGAC 340
C 361 ctgagagagagtcctcttccctcatcaccctgactcacaagaacaaagggcgagctct 420
D 399 CT-GGAGGAGGTGCTCTTTTCCCTCATCACTGACTCCACAAGAAACAAGGCGAGCTCT 281
C 421 catcactggggcttcaccactttcaggggtaag 453
D 260 CATCACI-GGGCTTACCACTTTTCAGGGTAAAG 249
RESULT 11
AX237229/c
LOCUS AX237229
DEFINITION Sequence 205 from Patent WO0164886.
ACCESSION AX237229
VERSION AX237229.1 GI:15796783
KEYWORDS human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 474)
AUTHORS Gaiger, A., Algate, P.A. and Mannion, J.
TITLE Compositions and methods for the detection, diagnosis and therapy
of hematological malignancies
JOURNAL Patent: WO 0164886-A 205 07-SEP-2001;
COPIXA CORPORATION (US)
FEATURES
Location/Qualifiers
1..474
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 165 a 70 c 111 g 126 t
ORIGIN 2 others

Query Match 71.8%; Score 334.4; DB 6; Length 474;
Best Local Similarity 99.2%; Pred. No. 2.9e-91;
Matches 357; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

C 1 atcaagaacacatagagttcggaacataatacttctcaccatcccccaaaactta 60
D 358 ATCAAGAAACATAGAGTT-GGGCAATATATCTTACCTTACCCATCCCAAACTTA 300
C 61 ctctactctctctctcattcattattttgggaataatcatcagaagatgtgttcagta 120
D 299 CTCTACTCATCTCATCTCATTATTTTGGGAATATCATCAGAGATGTGTCTGATGTA 240
C 121 agagattaaagaataaactttttgacccctcccaacaccccatccccagggtgacac 180
D 239 AGAGATTAAAGAAATTAAGCTTTTTCCTGCCAACACCCCATGCCAGGGTGTCAC 180
C 181 ctcccaatacaataacatccacagagatgaatttgccttttctgagtcggaatctgcc 240
D 179 CTCCCAATCAATAACATGCCAGAGAGTAAGTTGCCCTTTCTGATGCCGTAATTCGCC 120
C 241 atcatcttcccatcttccagttctcttccattgcaagtcacaatctgggtctcaggat 300
D 119 ATCATCTTCCCATCTTCCAGTCT-CTTTCATTTGCAAGTCACAACTTGGGTCTCAGGAT 61
C 201 ataccctcttagctctcgatcattgtcttccactgtgcactgagctgagaccttcgac 360
D 60 TATACCCGTCTTAGTCTCGATCACTTCTACCTTGTGCCACAGCTGAGCTGGACCTTCGAC 1

RESULT 12

AX237670/c	AX237670	474 bp	DNA	linear	PAI 26-SEP-2001
LOCUS	Sequence 645 from Patent WO0164886.				
DEFINITION	AX237670				
ACCESSION	AX237670				
VERSION	AX237670.1	GI:15797224			
KEYWORDS	human.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 474)				
TITLE	Gaiger, A., Algate, P.A. and Mannion, J.				
JOURNAL	Compositions and methods for the detection, diagnosis and therapy of hematological malignancies				
FEATURES	Patent: WO 0164886-A 546 07-SEP-2001;				
source	CORIXA CORPORATION (US)				
	Location/Qualifiers				
	1..474				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	165 a 70 c 111 g 125 t	2	others		
ORIGIN					
Query Match	71.8%; Score 334.4; DB 6; Length 174;				
Best Local Similarity	99.2%; Pred. No. 2.9e-91;				
Matches	357; Conservative 0; Mismatches 1; Indels 2; Gaps 2;				
QY	1 atcaagaacatagattcgggcaataatacttcactccaccacccaccccaatctta	60			
Db	358 ATCAAGAACAATAGAGTII-GGGCAATATATCTTCACTACCCATCCACCGAAATCTIA	300			
QY	61 ctctaccatctcatctcatttaattttggaaatcatcagaagaatgttctgttgaata	120			
Db	299 CTTACTATCTCTATCTCAATTAATTTTGGAAATCAATCAGAAATGTGTCTTISAGTA	240			
QY	121 agagattaaagaataaagctttttgacccctcgcaacaccccatgcccaaggttggatcac	180			
Db	239 AGAGATTAAGAAGAAATTAAGCTTTTGTATCCCTGCAACACCCCAATGCCAGGTCTAGTAC	180			
QY	181 cctccaatacaataacatccagaagaagatgaattgacctttctgataccataatggcc	240			
Db	179 CTTCCTAATCAATTAAGATGCCGGAAGAGTAAGTTGGCTTTCTGATGCCGTAATGTGCC	120			
QY	241 atcatctcccatcttcagctgccttttccattgcgaagtacacatctgggtcagagat	300			
Db	119 ATCATCTCCCATCTTCCAGTCT-CITTCATTCGAAGTCACAAATCTGGGCTTAGGGAT	61			
QY	301 tataccgcttttagtctgatacttacttctacttgcacactgaagcttgcac	360			
Db	60 TATACCGGCTTAGCTTCGATCAITGCTTACCTTGTGCCACTTGAGCTGGACCTTGCAC	1			
RESULT 13					
AX237272/c	AX237272	453 bp	DNA	linear	PAI 26-SEP-2001
LOCUS	Sequence 248 from Patent WO0164886.				
DEFINITION	AX237272				
ACCESSION	AX237272				
VERSION	AX237272.1	GI:15796826			
KEYWORDS	human.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 453)				
TITLE	Gaiger, A., Algate, P.A. and Mannion, J.				
JOURNAL	Compositions and methods for the detection, diagnosis and therapy of hematological malignancies				
FEATURES	Patent: WO 0164886-A 248 07-SEP-2001;				
source	CORIXA CORPORATION (US)				
	Location/Qualifiers				
	1..453				
	/organism="Homo sapiens"				

Query Match	32.5%	Score 151.4:	DB 10:	Length 6355:
Best Local Similarity	72.7%	Pred. No. 3.2e-35:		
Matches 293:	Conservative	0:	Mismatches 96:	Indels 14: Gaps 7:
Q1	54	tactcatctcatctctcatttaatttgggaagaatcatcagaagaagatcattc		

2865 GTCAGAGGAATCGGCTTGG---CTTCATACCGCATCCCC---GTGGTCATCCC 2813

[illegible]

364 ggaggagatgccttttgccctcatcactccaagaagaacaaaggcgacctgctcat
| | | | |
|| ||

2633 ccaaccgttcgaataaacatttggcccacacgagagcagcgcctgctcat
| | | | |
|| ||

2573 CACTGGGCTTCACCACTTCCAGGTAAAGGTGGATAGTGTTT 2532
SULT 15

17. Contains the gene for gamma-aminobutyric acid (GABA) B receptor (rhodopsin family, olfactory unkn) protein genes and one of the olfactory unkn

AL078530.1 GI:5051393
HTG: 7 transmembrane receptor; diubiquitin; GABA;
gamma-aminobutyric acid B receptor; GABA;
family: vomeronasal receptor type 1
family: vomeronasal receptor type 1

RENCE 1 (bases 1 to 134614)
THORS Younger, R.
ILE Direct Substratum

Where differences are found these are annotated as variations together with a note of the overlapping clones. During sequence assembly data is compared from overlapping clones. steve.sanger.ac.uk

been furnished according to sequence map criteria

```

Location/Qualifiers
1..154614
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="17"
/clone="573K1"
1..309
/note="match: GSS AB004433"
1286..2152
/note="putative CpG island"
1445..1574
/note="CpG island"

```

```

core
refseq_region
/feature="573K1.1" (gamma-aminobutyric acid (GABA) B
receptor, Ia)"
/evidence=not_experimental
1556..29807
/feature="573K1.1"
1533...1979
/feature="L1M3e_repeat: matches .34..325 of consensus"
join(2179..2260,2998..3141,3905..4090,5261..5281,
11768..6643,9621..9755,10229..10308,10575..10581,

```

```

/translation="MLLLLLPLFLRPLGAGAAQTPNTSEGCQIIHPHPWEGIRNG
LIIDQVKAINFLPVDYIEYVCREREYVGPVKRCLANGSTWMDTISRCVIRCSKS
YLTILENGKVLIGGLDGLARVDFRCDPFLHVGSRSTSCOGWSTPCHPCOVNR
TPHSERAVYIGALPMSGWPNGOACQPAEMALDVSNRDRLIDPYELKLIHHDS
CDQCATKIYELLYNDPIKILIMPGSSSVTLVAAEARMNLIIVLSVSSSPALSNR
QRFTIFERTHSATLHNPTRVKFEKMGWKKIATIQOTIEVFTSLDLDERVEKAGI
ETIFROSFSDPAPVKNLRKQAPILVGLFETEARKVCFCEYKTERLFGKKYWFLLI
GKXADNFKTIDPSINCEVMEATEVGEHTIELVILNLPANTRISINMISOEYKELT
KRLKRHEETGQFEAPLAYDAWALALUNKISGGGRSGRVLEDFNNYNOTIIDOI
YKLNSSSFEVGSVHPVFDASGRMAWILIEQLGGSGYKGIYGYDSTKDLKSKNTDI
WIGGSPADPOTVIKTRFLSQRFLTSVSLVSGIYI AWGJSENK
GPHLNHTAGC

```

NLIIAGUSGLAAVAPGLDGVHIGRSOFPEVCARUMLGLGSLGSGSMFT
 KINWNIHVIKKEEKKEWRKTLFPMKLYATVGLNGMDLITLAIWOIVDPHLRIIETI
 AREEKPEDIDVSILPQLEHCSKKMNTNLGIPYKGLGLLLGIFLAIYETKVSYTEKI
 HDHRAVGMAIYNVAVLCLITAPVTHILSSQDAAFASALAIYFSSYITLVVLFVFKM
 FELIIRGEWQSAQDITMKTSSTNNNEEKSRLLEKRELEKIIAEKEERYVSELHQ

Result	Query			DB	ID	Description
	No.	Score	Match Length			
C 1	399.8	85.8	833	21	AAC78081	Human cancer
C 2	397	85.2	777	19	AAV36265	Human diub
C 3	334.4	71.8	474	22	AAK54480	Human haem
C 4	334.4	71.8	474	22	AAK54921	Human haem
C 5	333.4	69.4	453	22	AAK54523	Human haem
C 6	234.5	50.3	666	23	AAS68568	DNA encoding
C 7	234.6	50.3	1129	23	AAS90737	DNA encoding
C 8	229.8	49.3	684	23	AAS71154	DNA encoding
C 9	229.8	49.3	684	23	AAS71927	DNA encoding

XX
XX
E2


```
Db 632 CTCACATCTCAATCTCAATTAATTTGGGAAATCATCAGAAATGATGTCGTHGAGIA 573
Oy 121 aagatataaagaataaagctttttgacccctgccaacaccccatccccaaggtggtcac 180
Db 572 AGAGATTAAAGAAATAGCTTTTGGCCCTGCGCAACACCCCAICCCAGGGTGGTCAC 513
Oy 181 cctccatacaataacatgccaaggaagtaagttgccccttctgataccgtaactgccc 240
Db 512 CCTCCATACATAATAGATGCGCAGGAAGTAAGTTGGCCCTTCTGATGCCCTAATCTGCC 453
Oy 241 atcattctcccatctccatctctcttccattgcaagtcacaatctgagatctcaagat 300
Db 452 ATATCTTCCCATCTCTCCAGTCT-CTTTCCATTCGACAGTCACAACTGGGGTCTCAGGGAT 394
Oy 301 tataccgctctgctcgatcattgttttccattgtgccaactgagctggagcccttcac 360
Db 393 TATACCGCTTATAGTCTCGATCATTTTTCACITGTGCCACATGAGCTTGGCAC 334
Oy 361 ctggagagagtgctcttttgcctcaccctgactccacaaagaacaaagggagctcct 420
Db 333 CT-GGAGGAGTGCTCTTTTGGCCCTCATCCTGACTCCACAAAGAACAAGGCGCAGCTCT 275
Oy 421 catcactggagcttaccacactttcaggagtaag 453
Db 274 CATCACT-GGGCTTCACACACTTTTCAGGGTAAGG 243

RESULT 3
AAK54480/c
ID AAK54480 standard: cDNA: 474 BP.
AC AAK54480:
XX
XX
DT 13-NOV-2001 (first entry)
XX
DE Human haematological malignancy-related antigen coding sequence #205.
XX
KW Human: cytostatic; vascular; gene therapy; vaccine; lymphoma;
KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
OS Homo sapiens.
XX
PR WO200164896-A2.
XX
PD 07-SEP-2001.
XX
PF 01-MAR-2001: 2001WO-US07272.
XX
PF 01-MAR-2000: 2000US-0186126.
PF 17-MAR-2000: 2000US-0150479.
PF 27-APR-2000: 2000US-0200545.
PF 28-APR-2000: 2000US-0200303.
PF 01-MAY-2000: 2000US-0200775.
PF 01-MAY-2000: 2000US-0200959.
PF 04-MAY-2000: 2000US-0202084.
PF 22-MAY-2000: 2000US-0206201.
PF 14-JUL-2000: 2000US-0216950.
PF 03-AUG-2000: 2000US-0222903.
PF 04-AUG-2000: 2000US-0223415.
PF 07-AUG-2000: 2000US-0223378.
XX
XX (CORI-) CORIXA CORP.
XX
PI Gaiger A, Algate PA, Mannion J.
XX
DR WPI: 2001-514842/56.
XX
PI Compositions and methods for the detection of hematological
PI malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
PI lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma -
XX
```

```
PS Claim 31: Pages 386-387: 1252pp: English.
XX
CC The present invention relates to compositions and methods for the
CC detection, diagnosis and therapy of haematological malignancies. The
CC present sequence is the coding sequence of a human haematological
CC malignancy related antigen. The methods of the present invention comprise
CC detecting the presence of haematological malignancy related antigen(s) in
CC a sample obtained from the patient (an increased level of the
CC polypeptide, compared to an unaffected individual, is indicative of an
CC increased risk). Haematological malignancies which can be treated using
CC the present invention are chronic lymphocytic leukaemia, lymphoma,
CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
CC cell non-Hodgkin's lymphoma.
XX
SQ Sequence 474 BP: 165 A: 70 C: 111 G: 126 T: 2 other:

Query: Vatch 71.8% Score 334.4: DB 22: Length 474:
Best Local Similarity 99.2%: Pred. No. 1.8e-98:
Matches 357: Conservative 0: Mismatches 1: Indels 2: Gaps 2:

Oy 1 atcaaaagaaacatagattcgggaataatacttctacccatcccccacaaatctta 60
Db 358 ATCAAGAAACAATAGATII-GGCAATATATCTTACCTACCCATCCACCAATCTTA 300
Oy 61 ctctactcatctcattctcatttgggaataatcagaagatgtgttcgtaagta 120
Db 229 CTCTACTCATCTCATTTCTCAITTAATTTGGGAAATCATCAGAGATGTGTCTGTGAGTA 240
Oy 121 agagattaaagaataaagctttttgacccctgccaacaccccatgccagggtggtcac 180
Db 239 AGAGATTAAAGAAATAAGCTTTTGTATCCTGCCACACCCCATGCCAGGGTGGTCAC 180
Oy 121 cctccatacaataacatgccaaggaagtaagttgccccttctgataccgtaactgccc 240
Db 119 CTCCTCAATACATAAATGATGCCAGGAAGTAAGTTGGCCCTTCTGATGCCCTAATCTGCC 120
Oy 241 atcattctcccatctccatctctcttccattgtgccaactgagctggagcccttcac 300
Db 119 ATCATCTTCCCATCTCTCCAGTCT-CTTTCCATTCGACAGTCACAACTGGGGTCTCAGGGAT 61
Oy 361 ctgagagagtgctcttttgcctcaccctgactccacaaagaacaaagggagctcct 360
Db 60 ATACCCGCTCTTGTGCTCGATCATTTTCACITGTGCCACATGAGCTTGGCAC 1

RESULT 4
AAK54921/c
ID AAK54921 standard: cDNA: 474 BP.
AC AAK54921:
XX
XX
DI 13-NOV-2001 (first entry)
XX
DE Human haematological malignancy-related antigen coding sequence #465.
XX
KW Human: cytostatic; vascular; gene therapy; vaccine; lymphoma;
KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
OS Homo sapiens.
XX
PR WO200164896-A2.
XX
PD 07-SEP-2001.
XX
PF 01-MAR-2001: 2001WO-US07272.
XX
PF 01-MAR-2000: 2000US-0186126.
PF 17-MAR-2000: 2000US-0150479.
PF 27-APR-2000: 2000US-0200545.
PF 28-APR-2000: 2000US-0200303.
PF 01-MAY-2000: 2000US-0200775.
PF 01-MAY-2000: 2000US-0200959.
PF 04-MAY-2000: 2000US-0202084.
PF 22-MAY-2000: 2000US-0206201.
PF 14-JUL-2000: 2000US-0216950.
PF 03-AUG-2000: 2000US-0222903.
PF 04-AUG-2000: 2000US-0223415.
PF 07-AUG-2000: 2000US-0223378.
XX
XX (CORI-) CORIXA CORP.
XX
PI Gaiger A, Algate PA, Mannion J.
XX
DR WPI: 2001-514842/56.
XX
PI Compositions and methods for the detection of hematological
PI malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
PI lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma -
XX
```

```

PR 01-MAY-2000: 2000US-0200999.
PR 04-MAY-2000: 2000US-0202084.
PR 22-MAY-2000: 2000US-0206201.
PR 14-JUL-2000: 2000US-0218950.
PR 03-AUG-2000: 2000US-0222903.
PR 04-AUG-2000: 2000US-0223416.
PR 07-AUG-2000: 2000US-0223378.
XX
XX (CORI-) CORIXA CORP.
XX
XX Gaiger A. Algate PA. Mannion J.
XX
XX WPI: 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological
XX malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
XX lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31: Page 499; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of hematological malignancies. The
XX present sequence is the coding sequence of a human hematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of hematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Hematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma.
XX
XX Sequence 474 BP: 155 A: 70 C: 111 G: 126 T: 2 other:

Query Match          71.8%   Score 334.4:   DB 22:   Length 474:
Best Local Similarity 99.2%   Pred. No. 1.9e-98:
Matches 357: Conservative 0: Mismatches 1: Indels 2: Gaps 2:

QY 1 atcaagaacacataaggttcgggcaataatactctaccatccaccacccaaatctta 60
DB 358 ATCAAGAACAACATAGAGTT-GGGCAATATACITTCATCCATCCCATCCCAATCTTA 300
QY 61 ctctactcatctcatctcatttaatttgggaaatcatcagaagatgtgtcgttgagta 120
DB 299 CTCTACTCATCTCAITCICATTAATTTTGGGAAATCATCAGAGATGTGTCTGTGAGTA 240
QY 121 agagattaaagaataaagctttttgacccctgccaacaccccccattgccagggtgtcac 180
DB 239 AGAGATTAAAGAATAAAGCTTTTTCATCCCTGCCAACACCCCATGCCAGGGTGTGCAC 180
QY 181 cctccaatacaataacatccaggaagataagttgcccctttctgtatccgtaactcac 240
DB 179 CTCTCAATACATATACATGCCAGAGAGTAAGTTCCTTTCTGATGCCGTAAATCTGCC 120
QY 241 atcatcttcccatcttccagctcccttccattgcaagtccacaaatctgggtctcagggat 300
DB 119 ATCATCTTCCCATCTTCCAGTCT-CITTCATTCGAAGTCACAAATCTGGGTCTCAGGAT 61
QY 301 tataccgctttagtctgcatcttgcctttcacttgcacactgagctggagcccttcgac 360
DB 60 TATACCCGCTTATAGTTCGATCATTTGCTTTCATCTTGTGCACCTGAGCTGACCTTCGCAC 1

RESULT 5
AAK54523/C
ID AAK54523 standard; cDNA: 453 BP.
XX
XX AAK54523:
XX
XX 13-NOV-2001 (first entry)
XX
XX Human hematological malignancy-related antigen coding sequence #249.

```

```

XX
XX Human: cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX hematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
XX Hcto sapiens.
XX
XX WO200154895-A2.
XX
XX 07-SEP-2001.
XX
XX 01-MAR-2001: 2001WO-US07272.
XX
XX 01-MAR-2000: 2000US-0186126.
XX 17-MAR-2000: 2000US-0190479.
XX 27-APR-2000: 2000US-0200545.
XX 23-APR-2000: 2000US-0200303.
XX 29-APR-2000: 2000US-0200779.
XX 01-MAY-2000: 2000US-0200999.
XX 04-MAY-2000: 2000US-0202084.
XX 22-MAY-2000: 2000US-0206201.
XX 14-JUL-2000: 2000US-0218950.
XX 03-AUG-2000: 2000US-0222903.
XX 04-AUG-2000: 2000US-0223416.
XX 07-AUG-2000: 2000US-0223378.
XX
XX (CORI-) CORIXA CORP.
XX
XX Gaiger A. Algate PA. Mannion J.
XX
XX WPI: 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological
XX malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
XX lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31: Page 399; 1252pp; English.
XX
XX The present invention relates to compositions and methods for the
XX detection, diagnosis and therapy of hematological malignancies. The
XX present sequence is the coding sequence of a human hematological
XX malignancy related antigen. The methods of the present invention comprise
XX detecting the presence of hematological malignancy related antigen(s) in
XX a sample obtained from the patient (an increased level of the
XX polypeptide, compared to an unaffected individual, is indicative of an
XX increased risk). Hematological malignancies which can be treated using
XX the present invention are chronic lymphocytic leukaemia, lymphoma,
XX follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX cell non-Hodgkin's lymphoma.
XX
XX Sequence 453 BP: 159 A: 67 C: 103 G: 123 T: 1 other:

Query Match          69.4%   Score 323.4:   DB 22:   Length 453:
Best Local Similarity 99.1%   Pred. No. 6.5e-95:
Matches 346: Conservative 0: Mismatches 1: Indels 2: Gaps 2:

QY 1 atcaagaacacataaggttcgggcaataatactctaccatccaccacccaaatctta 60
DB 347 ATCAAGAACAACATAGAGTT-GGGCAATATACITTCATCCATCCCATCCCAATCTTA 289
QY 61 ctctactcatctcatctcatttaatttgggaaatcatcagaagatgtgtcgttgagta 120
DB 299 CTCTACTCATCTCAITCICATTAATTTTGGGAAATCATCAGAGATGTGTCTGTGAGTA 229
QY 121 agagattaaagaataaagctttttgacccctgccaacaccccccattgccagggtgtcac 180
DB 228 AGAGATTAAAGAATAAAGCTTTTTCATCCCTGCCAACACCCCATGCCAGGGTGTGCAC 169
QY 161 cctccaatacaataacatccaggaagataagttgcccctttctgtatccgtaactcac 240
DB 159 CTCTCAATACATATACATGCCAGGAGAGTAAGTTCCTTTCTGATGCCGTAAATCTGCC 109

```

OY 241 atcatcttccatctccagatccttccattgcagatcacaatctgaggtctcagggat 300
|||||
DB 108 ATATCTTCCTCATCTTCCAGTCT-CTTTCATTTGCAAGTACAAATCTGGGCTTCAGGGAT 50
|||||
OY 301 tataccgccttagctcgcattgcttctcactctgacactgagctg 349
|||||
DB 49 TATACCGGCTTAGTCTGATCATCTTCTTCACTTTGGCAGTACGCTG 1
|||||
RESULT 6
AAS68568/c
ID AAS68568 standard: cDNA: 666 BP.
AC AAS68568;
XX
DI 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4372.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YI;
XX
DR WPI: 2001-639362/73.
DR P-PSDB; ABG04381.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 4372; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating

Query Match 50.3% Score 234.6; D9 23; Length 666;

Best Local Similarity 97.5%; Pred. No. 5.8e-66;
Matches 270; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
OY 177 tcacctccaataacaataacatgccagagaagagtaagttgccctttctgataccataatc 236
|||||
DB 666 TCACCTCCCAATACAATAAGATGCCAGGAAGAGTAGTTGCCCTTTCGTATCCGCTAATC 507
|||||
OY 237 tcccatcatcttcccatcttccagttctctcttccattgcaagtcacaatctgggtctcag 296
|||||
DB 606 TGCCATCATCTTCCCATCTTCCAGTCT-CTTTCATTTGCAAGTACAAATCTGGGCTCAG 548
|||||
OY 257 gaattataccctcttagtctcgaatcattgcttccacttgcactgagctggaccttc 356
|||||
DB 547 GGATTATACCCGCTTAGTCTCGATCATTTGCTTTTACATTGGCCACTGAGCTGGACCTTC 498
|||||
OY 357 gcacctggagagaggtgcctctttgcctcatcacctgactccacagaacaagagcgac 416
|||||
DB 487 GCACCT-GCAGGAGGTGGCTCTTTGCCCTCATCACCTGACTCCACAAGAAACAAAGGCGACG 429
|||||
OY 417 tctcatcactagggtctccaccactttcagggttaag 453
|||||
DB 429 TCTCATCACT-GGGCTTCACCACCTTTCAGGGTAAGG 393
|||||
RESULT 7
AAS90737/c
ID AAS90737 standard: cDNA: 1129 BP.
XX
AC AAS90737;
XX
DI 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #26541.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YI;
XX
DR WPI: 2001-639362/73.
DR P-PSDB; ABG26550.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 26541; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating

PF 30-MAR-2001: 2001WO-US08631.
XX
PP 31-MAR-2000: 2000US-0540217.
PR 23-AUG-2000: 2000US-0549157.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YI:
XX
XX WPI: 2001-639362/73.
XX P-PSDB: ABG07740.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity
XX
XX Claim 1: SEQ ID No 7731: 103pp: English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 684 BP: 233 A: 121 C: 200 G: 130 T: 0 other;
XX
XX Query Match 49.3%; Score 229.8; DB 23: Length 684;
XX Best Local Similarity 96.4%; Pred. No. 2.le-64;
XX Matches 267; Conservative 0; Mismatches 7; Indels 3; Gaps 3;
OY 177 tcaccctccaataacatgccaggagagtaagttgccccttctgataccgtaatc 236
DB 684 TCACCTCCCAATAAATAAGATGCCAGGAAGAGTAAGTTGCCCTTCTGATGCCGTAATC 625
OY 237 tgccatcatcttcccattctccagctctcttccattgcaagtccaactctgggtctcag 296
DB 624 TGGCATACTCTCCCACTCTCCAGTCT-CTTCCATTGCAAGTCAACTCTGGGCTCAG 566
OY 297 gattataccctgcttagtctgcatcattgcttctcacttgccactgagctgaccttc 356
DB 566 GGATTATACCCGCTTAGTCTCGATCATCTTCTTACCTTGTGCGCACTGAGCTGGACCTTC 506
OY 357 gcaactggggagggtgctcttctgctcatcactgactcactcacaagaagaaggaacac 416
DB 505 GCACCT-GGAGGAGGTGCTCTTCTCCCTCATCACTTGGCTGCCAAGAAACAAAGGGCAGC 447
OY 417 tctcatcactggggtctcaccacttccagggtgaag 453
DB 446 TCTTCATCACT-TGGGCTTCACCACTTTCAGGGTAAG 411
XX
XX RESULT 10
XX AAV36266/c
XX ID AAV36266 standard: cDNA; 496 BP.
XX

AC AAV36266:
XX
XX 29-SEP-1998 (first entry)
XX
XX Mouse diubiquitin A05F12 cDNA.
XX
XX Diubiquitin: mouse; dendritic cell; DC protein; cancer; diagnosis;
XX therapy: A05F12; ss.
XX
XX Mus sp.
XX
XX Key: Location/Qualifiers
XX CDS 8..495
XX /*tag= a
XX
XX W09823747-A2.
XX
XX 01-JUN-1998.
XX
XX 25-NOV-1997: 97WO-US20811.
XX
XX 11-DEC-1996: 96US-0763455.
XX 27-NOV-1996: 96US-0031806.
XX
XX (SCHE) SCHERING CORP.
XX
XX Barchereau J, Bates EEM, Caux C, De Saint-Vis BM;
XX Lecocq SJE:
XX
XX KFI: 3298-322727/28.
XX P-FSD: AAW50832.
XX
XX Dendritic cell genes - useful for developing products for diagnosis
XX and treatment of e.g. cancerous or degenerative conditions
XX
XX Example; Page 75; 92pp: English.
XX
XX This cDNA clone codes for a mouse diubiquitin protein (see AAW50832).
XX designated A05F12, which contains 2 ubiquitin domains and which
XX exhibits similarity to the monoclonal non-specific suppressor
XX factor beta produced by mouse and human T cells. The clone can
XX be expressed in a wide variety of host cells for the synthesis
XX of full-length A05F12 or polypeptide fragments. The novel DC
XX proteins A05F12 (see AAW50831-32), A07C03 (1g family) (see AAW50833-35),
XX and E2B02 (lysosome associated membrane protein (LAMP)-like) (see
XX AAW50935) are found in activated dendritic cells. The proteins play
XX a role in the regulation or development of haematopoietic cells,
XX e.g. lymphoid cells, which affect immunological responses. The
XX antigen presentation and the resulting effector functions. The
XX novel nucleic acids, polypeptides, and antibodies raised against
XX them, can be used in treatment of conditions associated with
XX abnormal physiology or development, including abnormal
XX proliferation, e.g. cancerous or degenerative conditions. The
XX products can also be used for detection, diagnosis or drug
XX screening.
XX
XX Sequence 496 BP: 147 A: 115 C: 125 G: 109 T: 0 other;
XX

Query Match 30.5%; Score 142; DB 19: Length 496;
XX Best Local Similarity 74.8%; Pred. No. 6.5e-36;
XX Matches 217; Conservative 0; Mismatches 70; Indels 3; Gaps 3;
OY 177 tcaccctccaataacatgccaggagagtaagttgccccttctgataccgtaatc 236
DB 496 TCATCCCCAGGAGTGTGTTGTGCAAGCAAGCAACTGCCACTCTTGAITGTGTGATC 437
OY 237 tgccatcatcttcccattctccagctctcttccattgcaagtccaactctgggtctcag 296
DB 436 AGCAATGATCTTTCATCTCCAG-CTCTCTCCGTTCGAATTCACAAGCTGCTCTTAC 378
OY 297 gattataccctgcttagtctgcatcattgcttctcacttgccactgagctgaccttc 356
XX

Db 377 GGATCAGAGGTCACACTCGATCATCTTTCACCTGGGCCACACGACGTCGATCTTC 318
 Oy 357 gcacctggaggaggtgctcttgcctcatcactgactccacaagaacaaggcagc 416
 Db 317 G-AACTCGGAGGAGGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 416
 Oy 417 tccctcatcactgggcttcaccatttcagggttaagtggtggtgcttt 465
 Db 258 TCITCATCCTT-GGGCTTCACCACTTCAGGCTAAGGTGGATAGTGGTTT 210

RESULT 11
 AAL33488
 ID AAL33488 standard: DNA: 51 BP.
 AC AAL33488;
 XX
 XX
 DT 24-JAN-2002 (first entry)
 XX
 DE Human SNP oligonucleotide #6696.
 XX
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinase; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200147944-A2.
 XX
 PD 05-JUL-2001.
 XX
 PF 28-DEC-2000: 2000WO-US35498.
 XX
 PR 28-DEC-1999: 99US-0173419.
 PR 27-DEC-2000: 2000US-0173419.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach M;
 XX
 DR WPI: 2001-465210/50.
 XX
 PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections.
 XX
 PS Claim 1: Page 3293; 4143pp; English.
 XX
 CC The present invention relates to oligonucleotides encoding polymorphic
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
 CC histones, kinases, colony stimulating factors, complement related
 CC proteins, cytochromes, kinases, cytokines, interferons, interleukins,
 CC G-protein coupled receptors and thioesterases. The present sequence is
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded
 CC by them may be used in the prevention, diagnosis and the treatment of
 CC diseases associated with inappropriate expression of the proteins listed
 CC above. Disorders that may be prevented, diagnosed and/or treated include
 CC multifactorial diseases with a genetic component, such as autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney, cancer
 CC leukaemia), diseases of the nervous system and an infection of pathogenic
 CC organisms.
 XX
 SQ Sequence 51 BP: 15 A; 13 C; 12 G; 11 T: 0 other:

Query Match 10.9%; Score 51; DB 22; Length 51;
 Best local similarity 100.0%; Pred. No. 8.6e-07;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 171 ggggggttcacccctcccaataacataacatgccaggagaagtaagtgcctt 221
 Db 1 ggggggttcacccctcccaataacataacatgccaggagaagtaagtgcctt 51

RESULT 12
 AAC42164/c
 ID AAC42164 standard: DNA: 934 BP.
 XX
 AC AAC42164;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 34526.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000: 2000EP-0301439.
 XX
 PR 25-FEB-1999: 99US-0121825.
 PR 05-MAR-1999: 99US-0133180.
 PR 05-MAR-1999: 99US-0133548.
 PR 23-MAY-1999: 99US-0125788.
 PR 25-MAR-1999: 99US-0126264.
 PR 29-MAR-1999: 99US-0126785.
 PR 01-APR-1999: 99US-0127462.
 PR 06-APR-1999: 99US-0128234.
 PR 09-APR-1999: 99US-0128714.
 PR 16-APR-1999: 99US-0129845.
 PR 19-APR-1999: 99US-0130077.
 PR 21-APR-1999: 99US-0130449.
 PR 23-APR-1999: 99US-0130510.
 PR 25-APR-1999: 99US-0130891.
 PR 30-APR-1999: 99US-0131449.
 PR 30-APR-1999: 99US-0132049.
 PR 01-MAY-1999: 99US-0132407.
 PR 05-MAY-1999: 99US-0132484.
 PR 06-MAY-1999: 99US-0132485.
 PR 06-MAY-1999: 99US-0132486.
 PR 07-MAY-1999: 99US-0132487.
 PR 11-MAY-1999: 99US-0132863.
 PR 14-MAY-1999: 99US-0134256.
 PR 14-MAY-1999: 99US-0134218.
 PR 14-MAY-1999: 99US-0134219.
 PR 14-MAY-1999: 99US-0134221.
 PR 18-MAY-1999: 99US-0134370.
 PR 19-MAY-1999: 99US-0134768.
 PR 20-MAY-1999: 99US-0134941.
 PR 21-MAY-1999: 99US-0135124.
 PR 21-MAY-1999: 99US-0135353.
 PR 25-MAY-1999: 99US-0135629.
 PR 26-MAY-1999: 99US-0136021.
 PR 26-MAY-1999: 99US-0136792.
 PR 01-JUN-1999: 99US-0137222.
 PR 03-JUN-1999: 99US-0137528.
 PR 04-JUN-1999: 99US-0137502.
 PR 07-JUN-1999: 99US-0137724.
 PR 08-JUN-1999: 99US-0138094.
 PR 10-JUN-1999: 99US-0138540.
 PR 10-JUN-1999: 99US-0138847.

```
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 17-JUN-1999: 99US-0139453.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 21-JUN-1999: 99US-0139763.
PR 22-JUN-1999: 99US-0139817.
PR 23-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 24-JUN-1999: 99US-0140354.
PR 28-JUN-1999: 99US-0140595.
PR 29-JUN-1999: 99US-0140823.
PR 30-JUN-1999: 99US-0140991.
PR 01-JUL-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 02-JUL-1999: 99US-0142154.
PR 06-JUL-1999: 99US-0142055.
PR 08-JUL-1999: 99US-0142190.
PR 09-JUL-1999: 99US-0142803.
PR 12-JUL-1999: 99US-0142920.
PR 13-JUL-1999: 99US-0142977.
PR 14-JUL-1999: 99US-0143542.
PR 15-JUL-1999: 99US-0143624.
PR 16-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 19-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144532.
PR 21-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 22-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145087.
PR 23-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 26-JUL-1999: 99US-0145224.
PR 27-JUL-1999: 99US-0145275.
PR 27-JUL-1999: 99US-0145913.
PR 27-JUL-1999: 99US-0145918.
PR 28-JUL-1999: 99US-0145919.
PR 02-AUG-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 03-AUG-1999: 99US-0146389.
PR 04-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 05-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 06-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 09-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147493.
PR 10-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
```

```
PR 11-AUG-1999: 99US-0148319.
PR 12-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148565.
PR 13-AUG-1999: 99US-0148684.
PR 15-AUG-1999: 99US-0149368.
PR 17-AUG-1999: 99US-0149175.
PR 18-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149829.
PR 23-AUG-1999: 99US-0149902.
PR 25-AUG-1999: 99US-0149930.
PR 26-AUG-1999: 99US-0150566.
PR 27-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 30-AUG-1999: 99US-0151080.
PR 31-AUG-1999: 99US-0151303.
PR 01-SEP-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 01-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 11-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 01-OCT-1999: 99US-0156596.
PR 03-OCT-1999: 99US-0157117.
PR 03-OCT-1999: 99US-0157753.
PR 03-OCT-1999: 99US-0157865.
PR 03-OCT-1999: 99US-0158029.
PR 11-OCT-1999: 99US-0158232.
PR 11-OCT-1999: 99US-0158369.
PR 11-OCT-1999: 99US-0159293.
PR 11-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 13-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 22-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 25-OCT-1999: 99US-0161359.
PR 25-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.
```

Query Match

Best Local Similarity 49.38; Score 50.2; DB 21; Length 934;

Matches 102; Conservative 18; Mismatches 86; Indels 1; Gaps 1;

```
07 181 ccccaataacaataatgcaggagaagtaagtgccttcttgatgcgtaattgcc 240
Db 301 ccccaataacaataatgcaggagaagtaagtgccttcttgatgcgtaattgcc 240
10-AUG-1999: 99US-0148171.
```



```
Db 1181 G-CGCTGCCAGCAAGATAGCCCTGCTGGTCGGAGGGAICCCCTTCCTTGCTCG 1123
QY 320 atcattgtttcactgtgccactgaagctgaccttcgcacctggagaggagt 372
Db 1122 ATTTTGTCTTCACATTAATCAACTGTGTCGGAGCTTTCACCTTCGAGGGIGAT 1070

RESULT 15
AAC24234/c
ID AAC382234 standard: DNA: 485 BP.
XX
AC AAC382234;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 20250.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000: 2000EF-0301439.
XX
PR 25-FEB-1999: 99US-0121825.
PR 05-MAR-1999: 99US-0123180.
PR 09-MAR-1999: 99US-0123548.
PR 23-MAR-1999: 99US-0125788.
PR 25-MAR-1999: 99US-0126264.
PR 29-MAR-1999: 99US-0126785.
PR 01-APR-1999: 99US-0127452.
PR 06-APR-1999: 99US-0128234.
PR 08-APR-1999: 99US-0128714.
PR 16-APR-1999: 99US-0129845.
PR 19-APR-1999: 99US-0130077.
PR 21-APR-1999: 99US-0130449.
PR 23-APR-1999: 99US-0130510.
PR 28-APR-1999: 99US-0130891.
PR 30-APR-1999: 99US-0131149.
PR 30-APR-1999: 99US-0132048.
PR 04-MAY-1999: 99US-0132407.
PR 05-MAY-1999: 99US-0132484.
PR 06-MAY-1999: 99US-0132485.
PR 06-MAY-1999: 99US-0132486.
PR 07-MAY-1999: 99US-0132487.
PR 11-MAY-1999: 99US-0132863.
PR 14-MAY-1999: 99US-0134356.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134221.
PR 18-MAY-1999: 99US-0134370.
PR 19-MAY-1999: 99US-0134768.
PR 20-MAY-1999: 99US-0134941.
PR 21-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138647.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 18-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145313.
PR 27-JUL-1999: 99US-0145318.
PR 27-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 02-AUG-1999: 99US-0146389.
PR 03-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 05-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 05-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 06-AUG-1999: 99US-0147493.
PR 09-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
PR 11-AUG-1999: 99US-0148319.
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 17:20:22 ; Search time 2552.63 Seconds
(without alignments)
1945.789 Million cell updates/sec

Title: US-09-880-107-3847
Perfect score: 368
Sequence: 1 gaagagacgtgtaagtgcg.....aataaatgtatctgaatc 368

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estopl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_gss:*
- 13: em_gss_hum:*
- 14: em_gss_inv:*
- 15: em_gss_pln:*
- 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	344.8	93.7	444	9	A1990990 wu35e04.x
C 2	343.2	93.3	388	9	AW584007 ia08e12.x
C 3	342.2	93.0	485	9	AW973307 EST395405
C 4	341.5	92.8	436	9	AA918966 o182e04.s
C 5	341.1	92.7	417	9	AA612829 na26h09.s
C 6	340.6	92.6	383	9	AA983320 oq56h04.s
C 7	340.2	92.4	399	10	BM090551 ig15a05.y
C 8	339.8	91.2	426	9	AA918970 o182e08.s
C 9	339.8	89.6	385	10	BM055886 id48b11.y
C 10	339.8	89.6	385	10	BM055886 id48b11.y
C 11	339.8	89.3	377	10	BI835329
C 12	325.4	88.4	395	9	BI835329
C 13	320.8	87.2	356	9	AA582941 ia08b11.y
C 14	319.8	86.9	718	10	AA551894 na13a05.s
C 15	319.2	86.7	370	9	BI836381
C 16	315.8	85.8	353	9	AA560251
C 17	314.8	85.5	318	9	AA845156

18	314.6	85.5	581	10	BI771478
C 19	309.6	84.1	373	9	AW471481
C 20	309.2	84.0	356	9	AW005713
C 21	294.4	79.9	337	9	AA835291
C 22	293.4	79.7	336	9	AA586834
C 23	293.6	78.7	329	9	AA845077
C 24	288.8	78.5	683	10	BI837747
C 25	285.6	77.9	321	9	AA569123
C 26	285.2	77.8	590	10	BI770923
C 27	285	77.4	329	9	AA921372
C 28	284.8	77.4	346	10	BG661253
C 29	283.8	77.1	317	9	AI362622
C 30	283.2	77.0	333	10	BM090732
C 31	278.2	75.6	314	10	BM090780
C 32	270.8	73.6	309	9	AI310323
C 33	257.2	72.6	308	9	AI459618
C 34	243.8	66.2	337	9	AI302919
C 35	243.6	66.2	316	10	BM090949
C 36	242.8	63.3	314	9	AA974970
C 37	230	52.5	405	10	BM432543
C 38	229.4	52.3	329	9	AA627501
C 39	228.4	52.1	365	10	BM431230
C 40	223	50.6	510	9	AI961220
C 41	215.8	58.6	371	9	AA835176
C 42	215	58.4	262	9	AA919095
C 43	207	56.2	250	9	AA844953
C 44	204.8	55.7	319	10	BM432834
C 45	204.4	55.5	583	10	BI837359

ALIGNMENTS

RESULT 1
A1990990/c
LOCUS
DEFINITION
wus35e04.x1 Soares-Dieckgraefe_colon_NHCD Homo sapiens CDNA clone
IMAGE:2522046.3' similar to gb:M1949 PANCRATIC SECRETORY TRYPSIN
INHIBITOR PRECURSOR (HUMAN);. mRNA sequence.
ACCESSION
A1990990
VERSION
A1990990.1 GI:5837887
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 444)
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -400P from Gibco
High quality sequence stop: 443.
FEATURES
Location/Qualifiers
Source
1..444
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2522046"
/tissue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Colon; Vector: pT730-Pac (Pharmacia) with a
modified polylinker; Site: 1: Not 1; Site: 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
IGTACCACTGAGTGGACGCGCCGCTTTTTHHHHHHHHHHH 3']
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraefe (Washington University, dieckim.wustl.edu); colonic mucosa represents a range of disease involvement from moderate to severe Crohn's disease; samples include both perforating (fistulas) and non-perforating samples. Library constructed by Bento Soares and M. Fatima Bonaldo.

```

BASE COUNT      118 a  105 c  103 g  118 t
ORIGIN

Query Match      93.7%  Score 344.8; DB 9; Length 444;
Best Local Similarity 96.8%; Pred. No. 1.3e-95;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 1 gaagagacgtgtaagtgcgggtcagattttcaactgacccctctggagcagaaacttcagcc 50
|||||
DB 384 GAAGAGACGTGTAAGTGGGGTGCAGCTTTCAACTGACCTCTGGAGCGCAGAACTTCAGCC 325
|||||

QY 61 atgaaggtaacagagcattttcttcagtgcttgcccttttaagtcattcttgtaac 120
|||||
DB 324 ATGAAGGTAAACAGGCATCTTCTCTCAGTGGCTTGGCCCTTGAGTCTATCIGIRAC 255
|||||

QY 121 actggagctgactccctggagagagggccaaatgttacaatgaacttaattgagacc 180
|||||
DB 254 ACTGGAGCTGACTCCCTGGGAAGAGAGGCCAAATGTTACAAATCAACTTAATGATGCACC 205
|||||

QY 181 agatatatgacctgctctggagactgatgaaataacttaccatgaatgaatgcgtgta 240
|||||
DB 204 AAGATATAGACCTGTCTGTGGGACGTGAGGAATACITATCCCAATGAAATCGGIGTA 145
|||||

QY 241 tgtttgaaggtcggaacacccgagactttctatctcattcaaaatctgggcttactga 300
|||||
DB 144 TGTITTTGAAAATCGGAACGCGAGACTTCTATCCICATTCATTCMAAAATCTGGGCTTGCIGA 85
|||||

QY 301 gaaccaaggtttgaaatccatccagctcagctcagccagagcc-----tattgttga 350
|||||
DB 84 GAACCAAGGTTTGAATCCCATCAGGTACCCGAGAGGCTGACTGGCCCTTATGTGTA 25
|||||

QY 351 taatgtatctgaaatc 358
|||||
DB 24 TAAATGATCTGAAATATC 7
|||||

```

```

RESULT 2
AW584007/c
LOCUS
DEFINITION
  gb:M11943 PANCREATIC SECRETORY TRYPsin INHIBITOR PRECURSOR (HUMAN);
  gb:X06342 Mouse mRNA for prostatic secretory glycoprotein (MOUSE);.
  mRNA sequence.
ACCESSION
  AW584007
VERSION
  AW584007.1 GI:7250981
KEYWORDS
  EST.
SOURCE
  human.
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 388)
  Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K.,
  Lemishka, I., Scaer, M., Brestelli, J., Gradwohl, G., Clifton, S.,
  Hillier, L., Marra, M., Pape, D., Wylie, I., Martin, J., Blais, A.,
  Schmitt, A., Theising, B., Ritter, E., Panko, I., Bennett, J., Cardenas
  M., Gibbons, M., McCann, R., Cole, R., Isagareishvili, R., Williams, I.,
  Jackson, Y., and Boxers, V.
  Endocrine Pancreas Consortium
  Unpublished (2000)
  Other ESTs: ia08e12.y1
  Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
  Endocrine Pancreas Consortium
  Harvard University, Howard Hughes Medical Institute
  Dept. of Molecular and Cellular Biology, 7 Divinity Ave. Cambridge,
  MA 02138

```

```

TITLE
  Unpublished (2000)
COMMENT
  Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
  Endocrine Pancreas Consortium
  Harvard University, Howard Hughes Medical Institute
  Dept. of Molecular and Cellular Biology, 7 Divinity Ave. Cambridge,
  MA 02138

```

```

Tel: 617-495-1812
Fax: 617-495-9557
Email: dmelcon@bioh.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brownjfas.harvard.edu)
Seq primer: 40up from Gibco.
Location/Qualifiers
  1..388
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone_lib="Human Pancreatic Islets"
  /tissue_type="Islets of Langerhans"
  /dev_stage="Adult"
  /lab_host="DH10B"
  /note="Organ: Pancreas; Vector: pSPORT1; Site: 1; Not 1:
  Site 2: Sal 1; Library constructed using SuperScript
  plasmid Library kit (Life technologies). cDNA made by
  oligo-dT priming. Size-selected by column fractionation;
  average insert size 1.08 kb. Primary library,
  unamplified."
BASE COUNT      105 a  92 c  84 g  106 t
ORIGIN

```

```

Query Match      93.3%  Score 343.2; DB 9; Length 388;
Best Local Similarity 96.6%; Pred. No. 3.9e-95;
Matches 365; Conservative 0; Mismatches 3; Indels 10; Gaps 1;

QY 1 gaagagacgtgtaagtgcgggtcagattttcaactgacccctctggagcagaaacttcagcc 50
|||||
DB 391 GAAGAGACGTGTAAGTGGGGTGCAGCTTTCAACTGACCTCTGGAGCGCAGAACTTCAGCC 322
|||||

QY 61 atgaaggtaacagagcattttcttcagtgcttgcccttttaagtcattcttgtaac 120
|||||
DB 321 ATGAAGGTAAACAGGCATCTTCTCTCAGTGGCTTGGCCCTTGAGTCTATCIGIRAC 262
|||||

QY 121 actggagctgactccctggagagagggccaaatgttacaatgaacttaattgagacc 180
|||||
DB 261 ACTGGAGCTGACTCCCTGGGAAGAGAGGCCAAATGTTACAAATCAACTTAATGATGCACC 202
|||||

QY 181 agatatatgacctgctctggagactgatgaaataacttaccatgaatgaatgcgtgta 240
|||||
DB 201 AAGATATAGACCTGTCTGTGGGACGTGAGGAATACITATCCCAATGAAATCGGIGTA 142
|||||

QY 241 tgtttgaaggtcggaacacccgagactttctatctcattcaaaatctgggcttactga 300
|||||
DB 141 TGTITTTGAAAATCGGAACGCGAGACTTCTATCCCTCATTCMAAAATATGGGCTTGCIGA 82
|||||

QY 301 gaaccaaggtttgaaatccatccagctcagctcagccagagcc-----tattgttga 350
|||||
DB 81 GAACCAAGGTTTGAATCCCATCAGGTACCCGAGAGGCTGACTGGCCCTTATGTGTA 22
|||||

QY 351 taatgtatctgaaatc 358
|||||
DB 21 TAAATGATCTGAAATATC 4
|||||

```

```

RESULT 3
AW973307/c
LOCUS
DEFINITION
  AW973307
ACCESSION
  AW973307
VERSION
  AW973307.1 GI:8163185
KEYWORDS
  EST.
SOURCE
  human.
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 485)
  Heide, P., Oi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt
  I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and

```

TITLE
Quackenbush, J.
Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray

JOURNAL
Unpublished (2000)

COMMENT
Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 317

```

seq primer: forward.
          location/Qualifiers
          1 485
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone_lib="MAGE"
          /rescues="MAGE"
          /not_vector="yes"
          119 a 115 9 130 t
BASE COUNT
ORIGIN

```

Query Match	93.0%	Score 342.2	DB 9	Length 485
Best Local Similarity	95.5%	Pred. No. 8.6e-95		
Matches 364	Conservative	0	Mismatches 3	Indels 10
				Gaps 1

Oy	1	gaagagacgttgtaagtcgcggtgcagtttccaactgacctctggacgcagaaattcagcc	60
Db	377	GAGGAGACGTGGTAAGTGC GGTCACTTTCAACTGACCITCGACCCAGAATTCAGCC	318
Oy	61	atgaaggtaaacgagcatctttcttcagtcgcttgcccctgttgagtctctatcggtaac	120
Db	317	ATGAAGGTAAACAGGCATCTTCTCAGTGCCCTTGGCCCTGTTGAGTCTATCTGGTAAC	258
Oy	121	actggagctgactccccgggaagagagccaaaatgttacaatgaacttaatggatgcc	180
Db	257	ACTGGAGCTGACTCCCTTGGGAAGAGAGGCCAATGTTACAATGAACCTAATGGAATGACC	198
Oy	181	aagatatagaccctgctdggagactgatggaataacttatcccaatgaatgcgtgta	240
Db	197	AAGATATATGACCCTGTCTGGGACTGATGGAATACTTATCCCAAIGAATGCGTGTA	138
Oy	241	tgtttgaaggtcggaacacccagactctctatccctcattcaaaaaatcggccctgtgta	300
Db	137	TGTTTTGAANAATCGGAACGCCAGACTTCTATCCTCATTCAAAAAICTGGCCCTTGTGTA	78
Oy	301	gaaccaaggtttgaaatcccatcaggctcacccgcgagcc-----tattgtttaa	350
Db	77	GAACCAAGGTTTTGAATCCCATCAGGTACCCGGGAGCCCTGACTGGCCCTATTGTTGAA	18
Oy	351	taaatgtatctgaatat	367
Db	17	IAAAIGTATCTGAATAT	1

RESULT	4
AA918966/c	
LOCUS	
DEFINITION	
AA918966 435 bp mRNA linear EST 17-APR-1999 ol82g04.sl NCI_CGAF_kid5 Homo sapiens cDNA clone IMAGE1546150 3. similar to gb:M11949 PANCREATIC SECRETORY TRYPSIN INHIBITOR PRECURSOR (HUMAN); mRNA sequence.	

ACCESSION	AA918965
VERSION	AA918966.1
KEYWORDS	GI:3058856
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 435) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index.
JOURNAL	Unpublished (1997)

Abstract

Contact: Robert Strausberg, Ph.D.
Email: cgabps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA library preparation: M. Bento Soares, Ph.D.
cDNA library arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:
www.bio.linnl.gov/bbrp/image/image.html
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 434.

U.
C.
C.
A.
C.
C.

```

1. .436
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1536150"
/clone_lib="NCL CGAP.kid5"
/tissue_type="2_pooled tumors (clear cell type)"
/lab_host="DH10b"
/note="Organ: Kidney; Vector: pT7T3D-Fac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'.
AATCGAGCAATTCGGCCGCAATATTTTTTTTTTTTTTTT 3'].
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo.
118 a 102 c 101 g 115 t

```

BASE CON
ORIGIN

Query Match	92.8%	Score 341.6	DB 9	Length 436
Best Seal Similarity	95.3%	Pred. No. 1.3e-94		
Matches 364	Conservative	0	Mismatches 4	Indels 10
			Caps 1	

Oy	1	gaagaagcgtggttaagtgcggtgcaggttttcaactgacctctggagcagaacttcagcc	50
Db	395	GAAGACAGCTGGTAAATGCGGTGCAGTITTTCAACTGACCTCTGGACGCGAAGCTTCAGCC	326
Oy	61	ataaaggtaaacagcgcattcttcttcagtcgcttgagccgttggtgctctatctgtgaac	120
Db	325	ATGAAGGTACAGSCATCTTCTTCTCAGTTCGCTTGGCCCTGTGAGTCTATCTGGTTAAC	256
Oy	121	actggagctgaactccctcgtggaagagagcgcaaatgtttacaaatgaacttaattggatgcacc	180
Db	265	ACTGGAGCTCACTCCCTGGGAAGAGAGGCCAANATGTTACAAATCAACTTATGAGATGCACC	206
Oy	161	agaatatatagccctctctgtggagactgatggaataacttatccccaatgaatcgctotta	240
Db	205	AAGATATATGACCCCTGCTGTGGAGCTGATGGAAATACTTATCCCAATGAATGCGGTTA	146
Oy	241	tattttgaaggtcgaaagccgcagactcttatcctcattccaaaaatctgggcttgcgtga	300
Db	145	TGTTTTGAAAAATCGGAACGCCAGACTTCTATGCTCATCAAAAATCTGGGCCCTTGCTGA	86
Oy	301	gaaccaaggttttgaaatcccatcaggttcacccgcgagcc-----tattgttgaa	350
Db	25	GAACCAAGGTTTTGAAATCCCATCAGGTCACCCGGTAGCTGAGCTTATGTTGAA	26
Oy	351	taaatatctcgaatc	368
Db	25	TAATGATCTGAATATC	8

6.
1-
2-
3-
4-
5-
6-

RESULT	F	417 bp	mrna	linear	EST 15-OCT-1997
LOCUS	AA512829				
DEFINITION	nm25h09.s1	NCI-CGAP_Col0	Homo sapiens cDNA clone IMAGE:114507		3'
ACCESSION	p2611949	PANCREATIC SECRETORY TRYPsin INHIBITOR			
	P2611949	(HUMAN);	mrna sequence.		
	AA512829				

```

VERSION      AA612829.1  GI:2453857
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 417)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL      Tumor Gene Index
COMMENT      Unpublished (1997)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.
              , Ph.D.
              CDNA Library Preparation: M. Bento Soares, Ph.D.
              CDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www.bio.llnl.gov/bbrp/image.html
              Insert length: 758 Std Error: 0.00
              Seq primer: -40m13 fwd. ET from Amersham
              High quality sequence stop: 395.
              Location/Qualifiers
                source
                  1..417
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:1145057"
                    /clone_lib="NCI-CGAP_Col0"
                    /tissue_type="colon tumor RER"
                    /lab_host="DH10B"
                    /note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
                    modified polylinker; 1st strand cDNA was prepared from
                    RER+ colon tumor, and was then primed with a Not I -
                    oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                    RI adaptors (Pharmacia), digested with Not I and cloned
                    into the Not I and Eco RI sites of the modified pT73
                    vector. Library is normalized. Library was constructed by
                    Bento Soares and M. Fatima Bonaldo (N-Soares4)."
              BASE COUNT  115 a  96 c  97 g  109 t
              ORIGIN
                1
                2
                3
                4
                5
                6
                7
                8
                9
                10
                11
                12
                13
                14
                15
                16
                17
                18
                19
                20
                21
                22
                23
                24
                25
                26
                27
                28
                29
                30
                31
                32
                33
                34
                35
                36
                37
                38
                39
                40
                41
                42
                43
                44
                45
                46
                47
                48
                49
                50
                51
                52
                53
                54
                55
                56
                57
                58
                59
                60
                61
                62
                63
                64
                65
                66
                67
                68
                69
                70
                71
                72
                73
                74
                75
                76
                77
                78
                79
                80
                81
                82
                83
                84
                85
                86
                87
                88
                89
                90
                91
                92
                93
                94
                95
                96
                97
                98
                99
                100
                101
                102
                103
                104
                105
                106
                107
                108
                109
                110
                111
                112
                113
                114
                115
                116
                117
                118
                119
                120
                121
                122
                123
                124
                125
                126
                127
                128
                129
                130
                131
                132
                133
                134
                135
                136
                137
                138
                139
                140
                141
                142
                143
                144
                145
                146
                147
                148
                149
                150
                151
                152
                153
                154
                155
                156
                157
                158
                159
                160
                161
                162
                163
                164
                165
                166
                167
                168
                169
                170
                171
                172
                173
                174
                175
                176
                177
                178
                179
                180
                181
                182
                183
                184
                185
                186
                187
                188
                189
                190
                191
                192
                193
                194
                195
                196
                197
                198
                199
                200
                201
                202
                203
                204
                205
                206
                207
                208
                209
                210
                211
                212
                213
                214
                215
                216
                217
                218
                219
                220
                221
                222
                223
                224
                225
                226
                227
                228
                229
                230
                231
                232
                233
                234
                235
                236
                237
                238
                239
                240
                241
                242
                243
                244
                245
                246
                247
                248
                249
                250
                251
                252
                253
                254
                255
                256
                257
                258
                259
                260
                261
                262
                263
                264
                265
                266
                267
                268
                269
                270
                271
                272
                273
                274
                275
                276
                277
                278
                279
                280
                281
                282
                283
                284
                285
                286
                287
                288
                289
                290
                291
                292
                293
                294
                295
                296
                297
                298
                299
                300
                301
                302
                303
                304
                305
                306
                307
                308
                309
                310
                311
                312
                313
                314
                315
                316
                317
                318
                319
                320
                321
                322
                323
                324
                325
                326
                327
                328
                329
                330
                331
                332
                333
                334
                335
                336
                337
                338
                339
                340
                341
                342
                343
                344
                345
                346
                347
                348
                349
                350
                351
                352
                353
                354
                355
                356
                357
                358
                359
                360
                361
                362
                363
                364
                365
                366
                367
                368
                369
                370
                371
                372
                373
                374
                375
                376
                377
                378
                379
                380
                381
                382
                383
                384
                385
                386
                387
                388
                389
                390
                391
                392
                393
                394
                395
                396
                397
                398
                399
                400
                401
                402
                403
                404
                405
                406
                407
                408
                409
                410
                411
                412
                413
                414
                415
                416
                417
                418
                419
                420
                421
                422
                423
                424
                425
                426
                427
                428
                429
                430
                431
                432
                433
                434
                435
                436
                437
                438
                439
                440
                441
                442
                443
                444
                445
                446
                447
                448
                449
                450
                451
                452
                453
                454
                455
                456
                457
                458
                459
                460
                461
                462
                463
                464
                465
                466
                467
                468
                469
                470
                471
                472
                473
                474
                475
                476
                477
                478
                479
                480
                481
                482
                483
                484
                485
                486
                487
                488
                489
                490
                491
                492
                493
                494
                495
                496
                497
                498
                499
                500
                501
                502
                503
                504
                505
                506
                507
                508
                509
                510
                511
                512
                513
                514
                515
                516
                517
                518
                519
                520
                521
                522
                523
                524
                525
                526
                527
                528
                529
                530
                531
                532
                533
                534
                535
                536
                537
                538
                539
                540
                541
                542
                543
                544
                545
                546
                547
                548
                549
                550
                551
                552
                553
                554
                555
                556
                557
                558
                559
                560
                561
                562
                563
                564
                565
                566
                567
                568
                569
                570
                571
                572
                573
                574
                575
                576
                577
                578
                579
                580
                581
                582
                583
                584
                585
                586
                587
                588
                589
                590
                591
                592
                593
                594
                595
                596
                597
                598
                599
                600
                601
                602
                603
                604
                605
                606
                607
                608
                609
                610
                611
                612
                613
                614
                615
                616
                617
                618
                619
                620
                621
                622
                623
                624
                625
                626
                627
                628
                629
                630
                631
                632
                633
                634
                635
                636
                637
                638
                639
                640
                641
                642
                643
                644
                645
                646
                647
                648
                649
                650
                651
                652
                653
                654
                655
                656
                657
                658
                659
                660
                661
                662
                663
                664
                665
                666
                667
                668
                669
                670
                671
                672
                673
                674
                675
                676
                677
                678
                679
                680
                681
                682
                683
                684
                685
                686
                687
                688
                689
                690
                691
                692
                693
                694
                695
                696
                697
                698
                699
                700
                701
                702
                703
                704
                705
                706
                707
                708
                709
                710
                711
                712
                713
                714
                715
                716
                717
                718
                719
                720
                721
                722
                723
                724
                725
                726
                727
                728
                729
                730
                731
                732
                733
                734
                735
                736
                737
                738
                739
                740
                741
                742
                743
                744
                745
                746
                747
                748
                749
                750
                751
                752
                753
                754
                755
                756
                757
                758
                759
                760
                761
                762
                763
                764
                765
                766
                767
                768
                769
                770
                771
                772
                773
                774
                775
                776
                777
                778
                779
                780
                781
                782
                783
                784
                785
                786
                787
                788
                789
                790
                791
                792
                793
                794
                795
                796
                797
                798
                799
                800
                801
                802
                803
                804
                805
                806
                807
                808
                809
                810
                811
                812
                813
                814
                815
                816
                817
                818
                819
                820
                821
                822
                823
                824
                825
                826
                827
                828
                829
                830
                831
                832
                833
                834
                835
                836
                837
                838
                839
                840
                841
                842
                843
                844
                845
                846
                847
                848
                849
                850
                851
                852
                853
                854
                855
                856
                857
                858
                859
                860
                861
                862
                863
                864
                865
                866
                867
                868
                869
                870
                871
                872
                873
                874
                875
                876
                877
                878
                879
                880
                881
                882
                883
                884
                885
                886
                887
                888
                889
                890
                891
                892
                893
                894
                895
                896
                897
                898
                899
                900
                901
                902
                903
                904
                905
                906
                907
                908
                909
                910
                911
                912
                913
                914
                915
                916
                917
                918
                919
                920
                921
                922
                923
                924
                925
                926
                927
                928
                929
                930
                931
                932
                933
                934
                935
                936
                937
                938
                939
                940
                941
                942
                943
                944
                945
                946
                947
                948
                949
                950
                951
                952
                953
                954
                955
                956
                957
                958
                959
                960
                961
                962
                963
                964
                965
                966
                967
                968
                969
                970
                971
                972
                973
                974
                975
                976
                977
                978
                979
                980
                981
                982
                983
                984
                985
                986
                987
                988
                989
                990
                991
                992
                993
                994
                995
                996
                997
                998
                999
                1000
                1001
                1002
                1003
                1004
                1005
                1006
                1007
                1008
                1009
                1010
                1011
                1012
                1013
                1014
                1015
                1016
                1017
                1018
                1019
                1020
                1021
                1022
                1023
                1024
                1025
                1026
                1027
                1028
                1029
                1030
                1031
                1032
                1033
                1034
                1035
                1036
                1037
                1038
                1039
                1040
                1041
                1042
                1043
                1044
                1045
                1046
                1047
                1048
                1049
                1050
                1051
                1052
                1053
                1054
                1055
                1056
                1057
                1058
                1059
                1060
                1061
                1062
                1063
                1064
                1065
                1066
                1067
                1068
                1069
                1070
                1071
                1072
                1073
                1074
                1075
                1076
                1077
                1078
                1079
                1080
                1081
                1082
                1083
                1084
                1085
                1086
                1087
                1088
                1089
                1090
                1091
                1092
                1093
                1094
                1095
                1096
                1097
                1098
                1099
                1100
                1101
                1102
                1103
                1104
                1105
                1106
                1107
                1108
                1109
                1110
                1111
                1112
                1113
                1114
                1115
                1116
                1117
                1118
                1119
                1120
                1121
                1122
                1123
                1124
                1125
                1126
                1127
                1128
                1129
                1130
                1131
                1132
                1133
                1134
                1135
                1136
                1137
                1138
                1139
                1140
                1141
                1142
                1143
                1144
                1145
                1146
                1147
                1148
                1149
                1150
                1151
                1152
                1153
                1154
                1155
                1156
                1157
                1158
                1159
                1160
                1161
                1162
                1163
                1164
                1165
                1166
                1167
                1168
                1169
                1170
                1171
                1172
                1173
                1174
                1175
                1176
                1177
                1178
                1179
                1180
                1181
                1182
                1183
                1184
                1185
                1186
                1187
                1188
                1189
                1190
                1191
                1192
                1193
                1194
                1195
                1196
                1197
                1198
                1199
                1200
                1201
                1202
                1203
                1204
                1205
                1206
                1207
                1208
                1209
                1210
                1211
                1212
                1213
                1214
                1215
                1216
                1217
                1218
                1219
                1220
                1221
                1222
                1223
                1224
                1225
                1226
                1227
                1228
                1229
                1230
                1231
                1232
                1233
                1234
                1235
                1236
                1237
                1238
                1239
                1240
                1241
                1242
                1243
                1244
                1245
                1246
                1247
                1248
                1249
                1250
                1251
                1252
                1253
                1254
                1255
                1256
                1257
                1258
                1259
                1260
                1261
                1262
                1263
                1264
                1265
                1266
                1267
                1268
                1269
                1270
                1271
                1272
                1273
                1274
                1275
                1276
                1277
                1278
                1279
                1280
                1281
                1282
                1283
                1284
                1285
                1286
                1287
                1288
                1289
                1290
                1291
                1292
                1293
                1294
                1295
                1296
                1297
                1298
                1299
                1300
                1301
                1302
                1303
                1304
                1305
                1306
                1307
                1308
                1309
                1310
                1311
                1312
                1313
                1314
                1315
                1316
                1317
                1318
                1319
                1320
                1321
                1322
                1323
                1324
                1325
                1326
                1327
                1328
                1329
                1330
                1331
                1332
                1333
                1334
                1335
                1336
                1337
                1338
                1339
                1340
                1341
                1342
                1343
                1344
                1345
                1346
                1347
                1348
                1349
                1350
                1351
                1352
                1353
                1354
                1355
                1356
                1357
                1358
                1359
                1360
                1361
                1362
                1363
                1364
                1365
                1366
                1367
                1368
                1369
                1370
                1371
                1372
                1373
                1374
                1375
                1376
                1377
                1378
                1379
                1380
                1381
                1382
                1383
                1384
                1385
                1386
                1387
                1388
                1389
                1390
                1391
                1392
                1393
                1394
                1395
                1396
                1397
                1398
                1399
                1400
                1401
                1402
                1403
                1404
                1405
                1406
                1407
                1408
                1409
                1410
                1411
                1412
                1413
                1414
                1415
                1416
                1417
                1418
                1419
                1420
                1421
                1422
                1423
                1424
                1425
                1426
                1427
                1428
                1429
                1430
                1431
                1432
                1433
                1434
                1435
                1436
                1437
                1438
                1439
                1440
                1441
                1442
                1443
                1444
                1445
                1446
                1447
                1448
                1449
                1450
                1451
                1452
                1453
                1454
                1455
                1456
                1457
                1458
                1459
                1460
                1461
                1462
                1463
                1464
                1465
                1466
                1467
                1468
                1469
                1470
                1471
                1472
                1473
                1474
                1475
                1476
                1477
                1478
                1479
                1480
                1481
                1482
                1483
                1484
                1485
                1486
                1487
                1488
                1489
                1490
                1491
                1492
                1493
                1494
                1495
                1496
                1497
                1498
                1499
                1500
                1501
                1502
                1503
                1504
                1505
                1506
                1507
                1508
                1509
                1510
                1511
                1512
                1513
                1514
                1515
                1516
                1517
                1518
                1519
                1520
                1521
                1522
                1523
                1524
                1525
                1526
                1527
                1528
                1529
                1530
                1531
                1532
                1533
                1534
                1535
                1536
                1537
                1538
                1539
                1540
                1541
                1542
                1543
                1544
                1545
                1546
                1547
                1548
                1549
                1550
                1551
                1552
                1553
                1554
                1555
                1556
                1557
                1558
                1559
                1560
                1561
                1562
                1563
                1564
                1565
                1566
                1567
                1568
                1569
                1570
                1571
                1572
                1573
                1574
                1575
                1576
                1577
                1578
                1579
                1580
                1581
                1582
                1583
                1584
                1585
                1586
                1587
                1588
                1589
                1590
                1591
                1592
                1593
                1594
                1595
                1596
                1597
                1598
                1599
                1600
                1601
                1602
                1603
                1604
                1605
                1606
                1607
                1608
                1609
                1610
                1611
                1612
                1613
                1614
                1615
                1616
                1617
                1618
                1619
                1620
                1621
                1622
                1623
                1624
                1625
                1626
                1627
                1628
                1629
                1630
                1631
                1632
                1633
                1634
                1635
                1636
                1637
                1638
                1639
                1640
                1641
                1642
                1643
                1644
                1645
                1646
                1647
                1648
                1649
                1650
                1651
                1652
                1653
                1654
                1655
                1656
                1657
                1658
                1659
                1660
                1661
                1662
                1663
                1664
                1665
                1666
                1667
                1668
                1669
                1670
                1671
                1672
                1673
                1674
                1675
                1676
                1677
                1678
                1679
                1680
                1681
                1682
                1683
                1684
                1685
                1686
                1687
                1688
                1689
                1690
                1691
                1692
                1693
                1694
                1695
                1696
                1697
                1698
                1699
                1700
                1701
                1702
                1703
                1704
                1705
                1706
                1707
                1708
                1709
                1710
                1711
                1712
                1713
                1714
                1715
                1716
                1717
                1718
                1719
                1720
                1721
                1722
                1723
                1724
                1725
                1726
                1727
                1728
                1729
                1730
                1731
                1732
                1733
                1734
                1735
                1736
                1737
                1738
                1739
                1740
                1741
                1742
                1743
                1744
                1745
                1746
                1747
                1748
                1749
                1750
                1751
                1752
                1753
                1754
                1755
                1756
                1757
                1758
                1759
                1760
                1761
                1762
                1763
                1764
                1765
                1766
                1767
                1768
                1769
                1770
                1771
                1772
                1773
                1774
                1775
                1776
                1777
                1778
                1779
                1780
                1781
                1782
                1783
                1784
                1785
                1786
                1787
                1788
                1789
                1790
                1791
                1792
                1793
                1794
                1795
                1796
                1797
                1798
                1799
                1800
                1801
                1802
                1803
                1804
                1805
                1806
                1807
                1808
                1809
                1810
                1811
                1812
                1813
                1814
                1815
                1816
                1817
                1818
                1819
                1820
                1821
                1822
                1823
                1
```

```

|||||
Db 203 GATATATACCCCTGTCTGTGGGACTGATGAATACTATATCCCAATGAATCGGTGTAATG 144
QY 243 tttaaaggctgaagacgcagactctctatccctcaatcaaaaatctgggcttgctgaaga 302
|||||
Db 143 TTTTGAATAATCGGAACGCCAGACTTCTATCTCTATTCATTCATCAAAATCTGGGCCCTGCTGAGA 84
QY 303 accaaggttttgaataatccctcaatcaagctcaacccaga-----ggcctattttgaataa 353
|||||
Db 83 ACCAAGGTTTGTGAATCCCATCAGGTACCGCGGAGCTGACTGGCCCTTATTTGTGAATAA 24
|||||
QY 354 atgtatctgaatctc 368
Db 23 ATGTATCTGAATATC 9
|||||

RESULT 7
BM090551 399 bp mRNA linear EST 20-NOV-2001
LOCUS
DEFINITION
ig15406.y1 Human Fetal Pancreas 1A Homo sapiens cDNA 5' similar to
SN:IPST1.HUMAN F00955 PANCREATIC SECRETORY TRYPsin INHIBITOR
PRECUSOR : mRNA sequence.
BM090551
ACCESSION
VERSION
BM090551.1 GI:17019517
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 399)
REFERENCE
MELTON D., BROWN J., KENTY G., PERMUTT A., LEE C., KAESTNER K.,
LEMISHKA J., SCEARCE M., BRESTELLI J., GRADWOHL G., CLIFTON S.,
HILLIER L., MARIA M., PAPE D., WYLIE J., MARTIN J., BLISTAIN A.,
SCHMITT A., THEISINGER B., RITTER E., KONKO I., BENNETT J., CARDENAS
M., GIBBONS M., MCCANN R., COLE R., TSAGAREISHVILI R., WILLIAMS T.,
JACKSON Y. and BOWERS Y.
Endocrine Pancreas Consortium
Unpublished (2000)
CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center for information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@imgate.wustl.edu)
Putative full length read
vector to vector length is.
FEATURES
Location/Qualifiers
1..399
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human Fetal Pancreas 1A"
/tissue_type="Fetal Pancreas (4 Pooled Donors, 18 - 20
weeks, Stratagene #738023)".
/derivative="Fetal Pancreas"
/note="vector: pBluescript SK(-); Site1: NotI; Site2:
XhoI; cDNA made by oligo-dT priming. Size selected by
CFP/MA SPIN-1000 (Clontech). Average insert size -1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110 E-mail: hinoue@imgate.wustl.edu Tel:
314-362-1916, Fax: 314-747-2692."
BASE COUNT
116 a 86 c 92 g 105 t
ORIGIN

```

```

Query: Match 92.4%; Score 340.2; DB 10; Length 399;
Best Local Similarity 96.5%; Pred. No. 3.3e-94; Mismatches 3; Indels 10; Gaps 1;
Matches 362; Conservative 0;
QY 4 gacacgtgtaagtcggtgcagatttttcaactgacctctggagcagagaacttcagccatg 63
|||||
Db 11 GAGACGTGGTAAGTGGGTCCAGTTTCAACTGACCTCTGGAGCGAGAACTTCAGCCATG 70
QY 64 aaggttaacgggactcttcttctcagtgctggcccttgaagcttatctagtgtaacact 123
|||||
Db 71 AAGGTAAACAGGCACTTCTCTCTCAGTGTCTGGCCCTGTGAGTCTATCTGGTAACACT 130
QY 124 gaaactgactcctcggaagaagccaaatgttaaatgaacttaaatgaatgcaccaa 183
|||||
Db 131 GGAGCTGACTCCCTGGGAAGAGAGGCCAATGTTACAACTTAATGGATGCACCAAG 190
QY 134 atatagacccctgtctgtggactgatgaaatacttatcccaatgaatgcgtgttatgt 243
|||||
Db 151 ATATATGACCCCTGCTGTGGGACTGATGGAATACTATCCCAATGAATCGGTGTTATGI 250
QY 243 tttaagggctcgaagaacgcagactctctcctcattcaaaaatctgggcttgctgaaga 303
|||||
Db 251 TTTGAAAATCGGAACGCCAGACTTCTATCTCTCAATTCATCAAAATCTGGGCCCTGCTGAGAA 310
QY 304 caaagttttgaaatcccatcaggtcacccgcagggc-----tattgttgaataa 353
|||||
Db 311 CCAAGGTTTTGAAATCCCATCAGTCACGTCGAGGCTGACTGGCCCTTATTTGTGAATAA 370
QY 354 atgtatctgaatctc 368
Db 371 ATGTATCTGAATATC 385
|||||

RESULT 8
AA918970 426 bp mRNA linear EST 10-JUN-1999
LOCUS
DEFINITION
ol82908.s1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1536158 3'
similar to gb:M11949 PANCREATIC SECRETORY TRYPsin INHIBITOR
PRECUSOR (HUMAN); mRNA sequence.
AA918970
ACCESSION
VERSION
AA918970.1 GI:3058860
KEYWORDS
ESI.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 426)
REFERENCE
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 623 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 409.
FEATURES
Location/Qualifiers
1..426
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI_CGAP_Kid5"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with

```


a modified polylinker; Site_1: Not 1; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-AATGAGAAATCGCGCCGCAATATTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 114 a 97 c 100 g 115 t

Query Match 91.2%; Score 335.8; DB 9; Length 426;
 Best Local Similarity 95.3%; Pred. No. 7.7e-93;
 Matches 361; Conservative 0; Mismatches 7; Indels 11; Gaps 1;

QY 1 gaagagagctgttaagtgggtgcagttttcaactgaacctctggagcagaaattcagcc 60
 |||||
 Db 386 GAAGAGAGCTGGTAAGTGGCGTGCAGTTTCACTGACCTCTGGAGCGCAATTCAGCC 327

QY 61 ataaagataacagagcatctttctcaatgccttgagcctgttgagttatctatgaac 120
 |||||
 Db 326 ATGAAGTAAACAGCATCTTCTTCAGTGGCTTGGCCCTGTGAGTCTATCTGTGAAC 257

QY 121 actggagctgcctccctggagagagagggccaaatgttacaataaacttaatgatcacc 180
 |||||
 Db 266 ACTGAGCTGACTCCCTGGAGAGAGGCCAAATGTTACAAATGAATTCATGATGCACC 207

QY 181 agatatataaccctgctgtggagctgatgaatactatccccaatgaatgcgtgtta 240
 |||||
 Db 206 AAGAATATAGCCCTGTCGTGGAGCTGATGGAATACTTATCCCAATGAATGCGTGTA 147

QY 241 tgtttgaagctggag 300
 |||||
 Db 146 TGTITTTGAAATCGGAACGCCAGACTTCTATCTTCATTCATTCATTCATTCATTCAT 87

QY 301 gaaccaaggtttgaataccatccatcaggtcacc-----gcgagggcctattgttga 349
 |||||
 Db 86 GAACCAAGGTTTGAATCCATCAGGTACCGCGGAGCCCTGACTGGCCCTTATTTGTA 27

QY 350 ataaatgtatctgaatc 368
 |||||
 Db 26 ATAAATGATCAGATATC 8

RESULT 9
 LOCUS BM055886 385 bp mRNA linear ESI 08-NOV-2001
 DEFINITION id84bl1.y1 Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1 Mus musculus cDNA 5' similar to SW:IPSI_HUMAN P00995 PANCREATIC SECRETORY TRYPsin INHIBITOR PRECURSOR 1; mRNA sequence.

ACCESSION BM055886
 VERSION BM055886.1 GI:15916497
 KEYWORDS ESI.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Chordata; Graniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurornathi; Muridae; Murinae; Mus. 1 (bases 1 to 385)

REFERENCE
 AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,I., Jackson,Y. and Bowers,Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Other_ESIs: id84bl1.x1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept. of Molecular and Cellular Biology, 7 Divinity Ave. Cambridge, MA 02138

Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@bioche.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Juliana Brown (brownjfas.harvard.edu).
 Location/Qualifiers
 1..385
 /organism="Mus musculus"
 /strain="ICR"
 /db_xref="taxon:10090"
 /clone_lib="Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1"
 /sex="Both for embryonic & newborn, male for adult and adult islet"
 /dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn, adult, mixed"
 /lab_host="DH10B"

/note="vector: pSPORT1; Site_1: Not 1; Site_2: Sal I; Five libraries representing E10.5/12.5 pancreatic bud, E16.5 pancreas, newborn pancreas, adult pancreas, and adult islets of Langerhans were separately constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA was made by oligo-dT priming and size-selected by column fractionation. Libraries were amplified once on solid support and plasmid DNA from each library was prepared and mixed in equal amounts. The mixed library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded mixed library plasmid DNA was mixed with 5 micrograms PCR product representing mixed library inserts and hybridized to an EcoT of 6. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 113 a 81 c 84 g 101 t

Query Match 89.6%; Score 329.8; DB 10; Length 385;
 Best Local Similarity 96.7%; Pred. No. 5.2e-91;
 Matches 351; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 16 gtcgggtcagattttcaactgaacctctggagcagaaacttcagccatgaagtaacagcc 75
 |||||
 Db 1 GTCGGTGCAGTTCACACIGACCTTCGACGACAGACTTCAGCCATGAAGTAACAGGC 60

QY 76 atctttctctcagtcgcttgagccctgttgagctctatctgtaaacactgagctgactcc 135
 |||||
 Db 81 ATCTTCTCTCAGTGGCTTGGCCCTGTGAGTCTATCTGTTAACTGAGCTGACTCC 120

QY 136 ctgggaagagagagccaaatgttacaataaacttaatgagtcacccaagatatatgacct 195
 |||||
 Db 121 CTGGGAGAGAGAGGCCAAATGTTACAAATGAAGTAAATGATGCACCAAGATATATGACCT 180

QY 196 atctgtggagctgatgaataactatcccaatgaatgcgtgttatgtttgaagtcgg 255
 |||||
 Db 191 GTCGTGGGACTGATGGAATACTTATCCCAATGAATGCGGTGTTATGTTTGAATTCGG 240

QY 256 taacgcagagactctatctctcatcaaaaactctggccttgcctgagaccaaagtttga 315
 |||||
 Db 241 AAACGCCAGACTTCTATCTTCATTCANAAATCTGGGCTTGTCTGAGAACCAAGTTTGA 300

QY 316 ataccatcaggtccacgcggagcc-----tattgttgataaataatgtatctgaat 365
 |||||
 Db 301 AATCCCATCAGGTACCGCGGAGGCTGACTGGCCCTTATTTGTAATTAATGATCTGAAT 360

QY 366 atc 368
 |||||
 Db 361 ATC 363

RESULT 10

```

B1835329      715 bp      mRNA      linear      EST 04-OCT-2001
LOCUS          50308504F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5224728 5'
DEFINITION     mRNA sequence.
ACCESSION      B1835329
VERSION        B1835329.1 GI:15946879
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 715)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: coapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: L14M1565 row: g column: 01
High quality sequence stop: 449.
Location/Qualifiers
1..715
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5224728"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-SPO16; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (In: Nitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT    251 a 166 c 190 g 108 t
ORIGIN

Query Match      89.6%; Score 329.8; DB 10; Length 715;
Best Local Similarity 96.7%; Pred. No. 6.7e-91;
Matches 351; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Q: 16 gtgcgtgcagttttcaactgacctctggacgcagaaacttcagccatgaaggtaacaggc 75
|||||
D: 1 GTGCGGTGCAGTTTTCAACIGACCTCTGGACGCCAGAACTTCAGGCATGAAGTTAACAGGC 60

Q: 76 atctttcttcagtcgttggccctgttgagttctatctgtatacactgagctgactcc 135
|||||
D: 61 ATCTTTCTCTCAGTGCCTTGCGCCGTGTGTGAGTCTATCTGTTAACTTGGACCTGACCTCC 120

Q: 136 ctgggaagagagcccaatgtttacataaacttaataatgagatgacccaagatataagacct 195
|||||
D: 121 CTGGGAGAGAGAGCCAAATGTTTACAACTGAACTTAATGATGACCAAGATATGACCTT 180

Q: 196 gctgtgggaactgatggaataacttatcccaatgaatgcgtgttatgtttgaaaggtcgg 255
|||||
D: 181 GTCTGTGGGACTGATGGAATACTATATCCCAATGAATGCGTGTTATGTTTTGAAAATCGG 240

Q: 256 aaacgcgaactctctctcattcaaaaactctggcccttgcctgagaccaggttttga 315
|||||
D: 241 AAACGCCAGACTTCTATCTCTCAATCAAAATCTGCGGCTCTCTGAGACCAAGGTTTGA 300

Q: 316 aatcccatcaggttcacccgcgagccc-----tattgttgaataaattgtatctgaat 365
|||||
D: 301 AATCCCATCAGGTTCACCGCGAGGCGCTGACTGCGCTTATGTTGTAATAAATGATCTGAA 350

```

Q: 346 atc 368
Db 361 ATC 363

RESULT 11

BG141336

LOCUS

DEFINITION

ia33c04.y3 Melton Human Islets H121 Homo sapiens cDNA 5' similar to

SW:IPST_HUMAN P00995 PANCREATIC SECRETORY TRYPSIN INHIBITOR

PRECURSOR: mRNA sequence.

ACCESSION

BG141336

VERSION

BG141336.1 GI:12644730

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 377)

Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Pape, D., Wylie, F., Martin, J., Blistain, A.,

Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas

, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, J.,

Jackson, Y., and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-9557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Juliana Brown

(brownj@fas.harvard.edu)

Seq primer: -40RP from Gibco.

Location/Qualifiers

1..377

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Melton Human Islets H121"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="TOP10"

/note="Organ: Pancreas; Vector: pZERO-2; Site_1: Not I;

Site_2: Xho I; Library constructed using SuperScript

Plasmid Library kit (Life Technologies). cDNA made by

oligo-dT priming. Xho I site destroyed during cloning.

Size-selected by column fractionation; average insert

size 1.59 kb. Primary library, unamplified."

BASE COUNT 112 a 81 c 83 g 101 t

ORIGIN

Query Match 89.3%; Score 328.8; DB 10; Length 377;

Best Local Similarity 96.7%; Pred. No. 1.1e-90;

Matches 350; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Q: 17 tgcgggtcagttttcaactgacctctggacgcagaaacttcagccatgaaggtaacaggca 76

|||||

D: 1 TCGGTGCAGTTTTCAACTGACCTCTGGACGCCAGAACTTCAGCCATGAAGTTAACAGGCA 60

Q: 77 tctttcttcagtcgcttggccctgttgagttctatctgttaacacactggaggtgactccc 135

|||||

D: 71 TCTTTCTTCAGTGCCTTGCCCTGTGTGAGTCTATCTGTTAACACTGGAGCTGACTCCC 120

Q: 137 tgggaagagagggcccaaatgtttacaataaacttaatgagatgcaccaagatatatgacccctg 195

```

|||||
Db 121 TGGAGAGAGGCCAAATGTTACAAAGCTTAAATGATGACCAACAGATATATGACCCCTG 180
Oy 197 tctgtggactgaggaataactatcccaatgaatcgtgttatgttttgaaggtcaga 256
Db 181 TCTGTGGACGTGAGAAATACITATCCCAATGNAATGCGTGTATGTTTTCAGAAATCGGA 240
Oy 257 aacccagacttctatcctcatcacaataatcggccttctctgagaccagaagttttgaa 315
Db 241 AACCCAGACATTCATCCATCAITCAAAAATCTGGGCCCTTGTGAGAACCAAGGTTTGAA 300
Oy 317 atcccatcagtcaccgcggagcc-----tattgttgaataaataatatctgaata 365
Db 301 ATCCCATCAGTCACCGGAGCGCTGACTGGCCCTTATTTGTAATAATATCTGAATA 360
Oy 367 tc 358
Db 361 TC 362

RESULT 12
AW582941
LOCUS
DEFINITION
  1a08b11.y1 Human Pancreatic Islets Homo sapiens cDNA, 5' similar to
  gb:M11949 PANC-REACTIC SECRETORY TRYPsin INHIBITOR PRECURSOR (HUMAN);
  gb:X05342 Mouse mRNA for prostatic secretory glycoprotein (MOUSE)).;
  mRNA sequence.
ACCESSION
  AW582941
VERSION
  AW582941.1 GI:7259985
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota: Chordata: Craniata: Vertebrata: Euteleostomi:
  Mammalia: Eutheria: Primates: Catarrhini: Hominidae: Homo.
  1 (bases 1 to 395)
REFERENCE
  Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
  Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
  Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
  Schmitt, A., Theising, B., Ritter, E., Panko, I., Bennett, J., Cardenas
  M., Gibbons, M., McGann, R., Cole, R., Tsagaris, R., Williams, J.,
  Jackson, Y., and Bowers, Y.
  Endocrine Pancreas Consortium
  Unpublished (2000)
  Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
  Endocrine Pancreas Consortium
  Harvard University, Howard Hughes Medical Institute
  Dept of Molecular and Cellular Biology, 7 Divinity Ave. Cambridge,
  MA 02138
  Tel: 617-495-1812
  Fax: 617-495-8557
  Email: dmelton@biohp.harvard.edu
  Library was constructed by Dr. Douglas Melton DNA sequencing by:
  Washington University; Genome Sequencing Center For information on
  obtaining a clone please contact: Juliana Brown
  (brownjfas.harvard.edu)
  Seq primer: -40RP from Gibco.
  Location/Qualifiers
    1..395
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="Human Pancreatic Islets"
    /tissue_type="Islets of Langerhans"
    /dev_stage="Adult"
    /lab_host="DH10B"
    /note="organ: Pancreas; Vector: pSPORT1; Site: 1; Rot 1;
    Site 2: Sal 1; Library constructed using SuperScript
    Plasmid Library kit (Life technologies). cDNA made by
    oligo-dt priming. Size-selected by column fractionation:
    average insert size 1.02 kb. Primary library,
    unamplified."
    105 a 87 c 95 g 107 t

BASE COUNT
ORIGIN

```

```

Query Match      88.4%; Score 325.4; DB 9; Length 395;
Best local Similarity 95.5%; Pred. NO. 1.2e-89;
Matches 360; Conservative 0; Mismatches 6; Indels 11; Gaps 2;

Oy 1 gaagagacgtggaagtcggtgcagttttcaactgaacctcttgagcgcagaacttcagcc 60
Db 19 GAAGAGACGTGGTAAAGTGGCGTGCAGTTCCTCAACTGACCTCTGAGCGCAGAACTTCAGCC 78
Oy 61 atcaaggttaacagagctatctttctctcagtcgcttgaccctgttgagttctat-ctggtaa 119
Db 79 ATGAGGTGAACAGGCACTCTTCTTCAGTGCCTTGCCCTGTGAGTCTATGGGGGAA 138
Oy 120 cactgagctgactccctcctgggaagagagccaaatgttacaatgaacttaatgagatcac 179
Db 132 CACTGAGCTGACTCCCTCGGAGAGAGAGCCAAATGTTACATGAACCTTAATGGATGCAC 198
Oy 190 caagatatatgacctctctgtgggactgatgaaalactatcccaatgaatgcattgtl 239
Db 193 CAAGAATATGACCTGTCTGTGGGACTGATGGAATACTTATCCCAATGAATGCGGT 258
Oy 240 atgttttgaaggtcgaagaaacccaaactctctcattcattcaaaaatctgggcttgcctg 299
Db 259 ATGTTTGAATAATCGGAAGCCAGAGCTTCTATCCTCATTCAAAAATCTGGGCCCTTGCTG 318
Oy 300 agaaccaaggttttgaatccccatcaggtccacgcgagcc-----tattattga 349
Db 312 AGAACCAAGGTTTGAATATCCCATCATGTCACCGGAGGCGCTGACCTGCTTATTGTCA 378
Oy 350 ataaatgtatctgaata 366
Db 353 ATAAATGATCTGAATA 395

RESULT 13
AW551894/C
LOCUS
DEFINITION
  NCI3a05.s1 NCI-CCAP_Co2 Homo sapiens cDNA clone IMAGE:1013360 3'
  similar to gb:M11949 PANC-REACTIC SECRETORY TRYPsin INHIBITOR
  PRECURSOR (HUMAN)).; mRNA sequence.
ACCESSION
  AW551894
VERSION
  AW551894.1 GI:2322146
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
  Mammalia: Eutheria: Primates: Catarrhini: Hominidae: Homo.
  1 (bases 1 to 356)
REFERENCE
  NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
  Emert-Buck, M.D., Ph.D.
  cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
  Ph.D.
  cDNA Library Arraying: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CCAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert length: 457 Std Error: 0.00
  Seq primer: -40ml3 fwd. ET from Amersham.
  Location/Qualifiers
    1..356
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="IMAGE:1013360"
    /tissue_type="tumor"
    /lab_host="SOLR (kanamycin resistant)"

```

/note="Organ: colon; Vector: Bluescript SK-; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
d1. Bulk colon villous adenoma. 5' adaptor sequence: 5'
GAATTCGGACGAG 3' 3' adaptor sequence: 5'
CTCAGTIIIIIIIIIIII 3' Average insert size: 1.1 kb.*
99 a 79 c 79 g 99 t

BASE COUNT

ORIGIN

Query Match 87.2%; Score 320.8; DB 9; Length 356;
Best Local Similarity 95.5%; Pred. No. 3e-88;
Matches 342; Conservative 0; Mismatches 2; Indels 10; Gaps 1;
QY 25 agtttcaactgaacttgagacgagaaacttcagccatgaaggtaaacagacatctttctt 84
Db 355 AGTIIITCAACTGACCTCTGGACGAGAACTTCAGCCATGAAGGTAAACAGGCATCTIIIT 297
QY 85 ctcaagtgccttggccctgttaagctctatcttgtaaacactggagctgacctccctgggaaga 144
Db 296 CTCAGTGCCTTGGCCCTCTTGAGCTATCTGGTAACTATGAGCTGACTCCCTGGGAAGA 237
QY 145 gaggccaatgttcaatgaacttaattggtatgagcaccacagatatatgacctgtctgtggg 204
Db 236 GAGGCCAATGTTTACAATGAACCTTAATGGATGCACCAAGAATATATGACCTGCTGTGGG 177
QY 205 actgaatgaatactatcccaatgaatgcgtgttatgttttcaaatgcgaacacccag 264
Db 176 ACTGATGAAATACCTATATCCCAATGAATGGGTGTTATGTTTGAATCGGAACGCCAG 117
QY 265 actttatctctcattcaaaaatctggccctgtgtgagaacacaggttttgaatcccatc 324
Db 116 ACTTCTATCTCTCATTCAAAAATCTGGCCCTGTCTGAGAACCAAGGTTTGAATCCCATC 57
QY 325 aggtcacccgagacc-----tatgttgaataaaatgtatctgaatatc 368
Db 56 AGGTACCCGCGAGCCCTGACTGGCCCTTATTGTTGAATAAATGATCTGAATAIC 3

RESULT 14
BI836381 718 bp mRNA linear EST 04-OCT-2001
LOCUS 603083093F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5222134 5',
mRNA sequence.
ACCESSION BI836381
VERSION BI836381.1 GI:15947931
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 718)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1558 row: j column: 23
High quality sequence stop: 426.
Location/Qualifiers
1..718
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5222134"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:

FEATURES

source

pcMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb. Insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library.*
239 a 156 c 198 g 115 t

BASE COUNT

ORIGIN

Query Match 86.9%; Score 319.8; DB 10; Length 718;
Best Local Similarity 95.6%; Pred. No. 8.2e-88;
Matches 341; Conservative 0; Mismatches 2; Indels 10; Gaps 1;
QY 26 gtttcaactgaacttgagacgagaaacttcagccatgaaggtaaacagacatctttctt 85
Db 1 GTIIITCAACTGACCTCTGGACGAGAACTTCAGCCATGAAGGTAAACAGGCATCTIIITC 60
QY 35 tcaagtgccttggccctgttgaagctctatcttgtaaacactggagctgacctccctgggaaga 145
Db 51 ICAGTGCCTTGGCCCTTTCAGTCTATCTGTAACACTGAGCTGACTCCCTGGGAAGAG 120
QY 116 agaccaaatgttacaatgaacttaattggtatgagcaccacagatatatgacctgtcttgaga 205
Db 121 AGCCCAATGTTTACAATGAACCTTAATGGATGCACCAAGAATATATGACCTGCTGTGGGA 180
QY 206 ctgaatgaatactatcccaatgaatgcgtgttatgttttgaagctggaacacccaga 265
Db 151 CTGATGAAATACCTATATCCCAATGAATGGGTGTTATGTTTGAATCGGAACGCCACA 240
QY 266 ctctatctctcattcaaaaatctggccctgtgtgagaacacaggttttgaatcccatca 325
Db 241 CTCTATCTCTCATTCAAAAATCTGGCCCTGTCTGAGAACCAAGGTTTGAATCCCATCA 300
QY 326 ggtcacccgagacc-----tatgttgaataaaatgtatctgaatatc 368
Db 301 GGTACCCGCGAGCCCTGACTGGCCCTTATTGTTGAATAAATGATCTGAATAIC 353

RESULT 15

BI650251

LOCUS

DEFINITION

370 bp mRNA linear EST 18-DEC-1999
we58h12.xl Soares_Dieckgraeft_colon_NHCD Homo sapiens cDNA clone
IMAGE:2346311 3' similar to gb:M11949 PANCREATIC SECRETORY TRYPSIN
INHIBITOR PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 370)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (<http://image.llnl.gov>) for further information.
Trace considered overall poor quality
Insert Length: 550 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1..370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2346311"
/clone_lib="Soares_Dieckgraeft_colon_NHCD"

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

```
/issue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: p173D-Pac (Pharmacia) with a
modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I... oligo(dT) primer (5'-
IGTACCAACATGAGAGGAGCGCGCGCIIIIIIIIIIIIIIIIIIII 3').
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p173 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraefe (Washington University,
St. Louis, MO; dieckim.wustl.edu); colonic mucosa represents a range of
disease involvement from moderate to severe Crohn's
disease; samples include both perforating (fistulas) and
non-perforating samples. Library constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT      105 a      98 c      78 g      98 t
ORIGIN
```

```
Query Match      86.7%: Score 319.2: DB 9: Length 370:
Best Local Similarity 93.8%: Pred. No. 9.5e-88:
Matches 347: Conservative 0: Mismatches 13: Indels 10: Gaps 1:

Oy 3 agagacgtgataagtcagatgcattttcaactgaacctctggagcagcaacttcagccat 62
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 370 AGAGACGTGTAGTGGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 311
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Oy 63 gaaggtaacaggcattcttcttcagtcgcttgccctgttgaagtcctatctagttaacac 122
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 310 GAAGGTAACAGGCACTCTTATCAGTGGCTTGGCCCTGTTGAGTCTATCTGGTAACAC 251
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Oy 123 tgaagctgaactccctgagaaagagaccataatttacaatgaacttaattgagaccacaa 182
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 250 TGAAGCTGACTCCCTGGGAGAGAGAGGCCAAATGTTGCAATGAACITTAATGGATGCCACAA 191
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Oy 183 gatatagaccctgtctgtgggactgatgaaataacttaccatgaatgagctgttatg 242
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 190 GATATATGACCCGTCGTGCGGAATGATGAATAACTTATCCCAATGAATGCGGTGTTATG 131
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Oy 243 ttltgaaggctggaagcagcagacttctatcctcattcaaaaaatctgagcccttgctgaga 302
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130 TTTTGAATAATCGGAACGCCAGCGCTTCTATCCTCATTCATAAAAAATTTGGGCTTGTGTAGA 71
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Oy 303 accaaggttttgaatcccatcaggtcaccgcagagcc-----tatgtttgaata 352
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 70 ACCAAGGTTTGAATAATCCCAATCAGGTCTCCGGAGGCGCTTATGTTGAATA 11
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Oy 353 aatgtatctg 362
    ||||||||||
Db 10 AATGATATCTG 1
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 19:27:19 ; Search time 394.41 Seconds
(without alignments)
1967.976 Million cell updates/sec

Title: US-09-880-107-3847

Perfect score: 368
Sequence: 1 gaagagacgtgtaagtgcg.....aataaatgtatcgaatc 368

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1333638 seqs, 1054605264 residues

Total number of hits satisfying chosen parameters: 2667276

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending_Patents_NA_Nev.*

1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	344.8	93.7	603	US-10-162-749-16	Sequence 16, Appl
2	344.8	93.7	607	US-10-162-749-57	Sequence 57, Appl
3	210.8	57.3	214	US-60-382-445-457	Sequence 457, App
4	88.8	24.1	180	US-09-454-226A-557	Sequence 557, App
5	45	12.2	421	US-60-365-384-494	Sequence 494, App
6	45	12.2	652	US-60-365-384-156	Sequence 156, App
7	42.2	11.5	323	PCT-US01-43704-206	Sequence 206, App
8	42.2	11.5	323	PCT-US01-43704-2330	Sequence 2330, App
9	42.2	11.5	398	US-09-671-483-2	Sequence 2, Appl
10	42.2	11.5	433	US-10-106-698-1897	Sequence 1897, App
11	40.6	11.0	3528	PCT-US02-10824-19	Sequence 19, Appl
12	40.6	11.0	3532	US-09-926-820-7	Sequence 7, Appl
13	40.4	11.0	156	US-09-926-820-12	Sequence 12, Appl
14	39.2	10.7	562	US-09-629-469A-8198	Sequence 8198, App
15	39.2	10.7	1022	US-09-629-469A-11819	Sequence 11819, A
16	38.2	10.4	180	US-09-671-483-1	Sequence 1, Appl
17	38	10.3	338	PCT-US02-10421-1433	Sequence 1433, App
18	38	10.3	338	US-10-112-699-1433	Sequence 1433, App
19	36.6	9.9	365	PCT-US02-10421-1457	Sequence 1457, App
20	36.6	9.9	365	US-10-112-699-1457	Sequence 1457, App
21	36.6	9.9	394	PCT-US02-10421-1212	Sequence 1212, App
22	36.6	9.9	394	US-10-112-699-1212	Sequence 1212, App
23	36.6	9.9	395	PCT-US02-10421-2209	Sequence 2209, App
24	36.6	9.9	395	US-10-112-699-2209	Sequence 2209, App
25	36.6	9.9	426	PCT-US02-10421-1089	Sequence 1089, App

26	36.6	9.9	426	7	US-10-112-699-1089	Sequence 1089, Ap
27	36.6	9.9	437	1	PCT-US02-10421-2071	Sequence 2071, Ap
28	36.6	9.9	437	1	US-10-112-699-2071	Sequence 2071, Ap
29	36.6	9.9	439	1	PCT-US02-10421-1527	Sequence 1527, Ap
30	36.6	9.9	439	1	PCT-US02-10421-2125	Sequence 2125, Ap
31	36.6	9.9	439	7	US-10-112-699-1527	Sequence 1527, Ap
32	36.6	9.9	439	7	US-10-112-699-2125	Sequence 2125, Ap
33	36.6	9.9	447	1	PCT-US02-10421-2349	Sequence 2349, Ap
34	36.6	9.9	447	7	US-10-112-699-2349	Sequence 2349, Ap
35	36.6	9.9	691	7	US-10-162-749-26	Sequence 26, Appl
36	36.6	9.9	695	7	US-10-143-788-50	Sequence 50, Appl
37	36.6	9.9	695	7	US-10-162-749-52	Sequence 52, Appl
38	36.6	9.9	700	7	US-10-143-788-53	Sequence 53, Appl
39	36.6	9.9	700	7	US-10-162-749-67	Sequence 67, Appl
40	36.6	9.9	746	7	US-10-162-749-11	Sequence 11, Appl
41	36.2	9.8	195	5	US-09-442-366A-1076	Sequence 1076, Ap
42	36.2	9.8	305	1	PCT-US02-10421-1244	Sequence 1244, Ap
43	36.2	9.8	305	7	US-10-112-699-1244	Sequence 1244, Ap
44	35	9.5	439	1	PCT-US02-10421-2521	Sequence 2521, Ap
45	35	9.5	439	7	US-10-112-699-2521	Sequence 2521, Ap

ALIGNMENTS

RESULT 1
US-10-162-749-16
: Sequence 16, Application US/10162749
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PTZ67C1N
: CURRENT APPLICATION NUMBER: US/10/162,749
: CURRENT FILING DATE: 2002-06-06
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 184
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 16
: LENGTH: 603
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (584)
: OTHER INFORMATION: n equals a,t,g, or c
US-10-162-749-16

Query Match	93.7%	Score	344.8	DB 7:	Length	603:
Best Local Similarity	96.8%	Pred. No.	5.5e-100;			
Matches	366:	Conservative	0:	Mismatches	2:	Indels
					10:	Gaps
Oy	1	gaagagacgtgtaagtgcggtgcagttttcaactgaccttgacgcagaaacttcagcc	60			
Db	177	gaagagacgtgtaagtgcggtgcagttttcaactgaccttgacgcagaaacttcagcc	236			
Oy	51	atgaaggttaacaggcattcttctcagtcgcttgccctgtgtgagtcgtatctgtgtaac	120			
Db	237	atgaaggttaacaggcattcttctcagtcgcttgccctgtgtgagtcgtatctgtgtaac	296			
Oy	121	actgagctgactccctgggaagagagccaaatgttacaatgaacttaattggtgacc	180			
Db	297	actgagctgactccctgggaagagagccaaatgttacaatgaacttaattggtgacc	356			
Oy	181	aagatatgacctgtctgtggaactgagtgaaataacttaccatgaatgcgtgtta	240			
Db	357	aagatatgacctgtctgtggaactgagtgaaataacttaccatgaatgcgtgtta	416			
Oy	261	tggttgaagtcgaaacccagacacttctatctcattcaaaaatctgggctgtgta	300			
Db	417	tggttgaagtcgaaacccagacacttctatctcattcaaaaatctgggctgtgta	476			
Oy	301	gaaccaagttttgaaatccctcaggtcacccgcgagcc-----tattgttgaa	350			

Db 477 gaaccagggtttgaaatcccatcaggtcaccgcgaggcctgactgagccttattgttgaa 536
Qy 351 taaatgtatctgaatc 368
Db 537 taaatgtatctgaatc 554

RESULT 2
US-10-162-749-57/c
: Sequence 57, Application US/10162749
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PTZ67CIN
: CURRENT APPLICATION NUMBER: US/10/162,749
: CURRENT FILING DATE: 2002-06-06
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 184
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 57
: LENGTH: 607
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (23)
: OTHER INFORMATION: n equals a,t,g, or c
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (25)
: OTHER INFORMATION: n equals a,t,g, or c
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (35)
: OTHER INFORMATION: n equals a,t,g, or c
US-10-162-749-57

Query Match 93.7% Score 344.8; DB 7; Length 607;
Best Local Similarity 96.8%; Pred. No. 5,5e-100;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Qy 1 gaagagacgtggaagtcgagtcagttttcaactgacctctggaacagaaactcagcc 60
Db 429 GAAGAGACGTGTGAAGTCGGGTGCAGTTTTCACCTGACCTTGGACGCAGAACTTCAGCC 370
Qy 61 atgaagtaacagacatcttcttctcagtcgcttgcctctgttgagtcctatctggtaac 120
Db 369 ATGAAGGTAACAGGACATCTTCTTCTCAGTGCCTTGCCCTGTTCAGTCTAICTGTTAAC 310
Qy 121 actgagctgactcctcctgggaagagagagccaaatgttacaatgaacttaattggtgacc 180
Db 309 ACTGGAGCTGACTCCTCGGAAGAGAGAGGCCAAATGTTACAAATGAACITTAATGGATGCACC 250
Qy 181 aagatatagcctctgtctggagactgagaaacttatcccaatgaatgcgtgta 240
Db 249 AAGATATATGACCCCTGTCTGTGGAGCTGATGGAATATCTTATCCCAATGAATGCGTGA 190
Qy 241 tgtttgaaggtcggaacgcagactctctatcctcattcaaaaatctgggacctgtga 300
Db 189 TGTTTGAAAATCGGAACGCCAGCAGACTTCTATCCTCATTAATAAATCTGGGCTTGTCTGA 130
Qy 301 gaaccaaggttttgaatcccatcaggtcaccgcgagcc-----tatgttga 350
Db 129 GAACCAAGGTTTGAATCCCATCAGGTACAGGTACCGCCGAGGCTGACTGGCCTTATTTGTA 70
Qy 351 taaatgtatctgaatc 368
Db 69 TAAATGATCTGAATATC 52

RESULT 3

US-60-382-445-457
: Sequence 457, Application US/60382445
: GENERAL INFORMATION:
: APPLICANT: Algate, Paul A.
: APPLICANT: Mannion, Jane
: APPLICANT: Gaiger, Alexander
: APPLICANT: Gordon, Brian
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: FILE REFERENCE: 210121.573PI
: CURRENT APPLICATION NUMBER: US/60/382,445
: CURRENT FILING DATE: 2002-05-20
: NUMBER OF SEQ ID NOS: 2900
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 457
: LENGTH: 214
: TYPE: DNA
: ORGANISM: Homo sapiens
US-60-382-445-457

Query Match 57.3% Score 210.8; DB 8; Length 214;
Best Local Similarity 99.1%; Pred. No. 2.4e-57;
Matches 212; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 125 gagctgactcctcctgggaagagagagccaaatgttacaatgaacttaattggtgaccacaaga 184
Db 1 gagctgactcctcctgggaagagagagccaaatgttacaatgaacttaattggtgaccacaaga 60
Qy 125 tatatgacctgtctgtggagactgatggaatactatcccaatgaatgcgtgtatgt 244
Db 61 tatatgacctgtctgtggagactgatggaatactatcccaatgaatgcgtgtatgt 120
Qy 245 ttgaaggtcggaacgcagactctctcctcattcaaaaatctgggacctgtgagaac 304
Db 121 ttgaaatcggaacgcagactctctcctcattcaaaaatctgggacctgtgagaac 180
Qy 305 taaggtttgaaatcccatcaggtcaccgcgaggg 338
Db 191 taaggtttgaaatcccatcaggtcaccgcgaggg 214

RESULT 4
US-09-454-226A-557
: Sequence 557, Application US/09454226A
: GENERAL INFORMATION:
: APPLICANT: Chenchik, Alex
: APPLICANT: Lukashev, Matvey
: TITLE OF INVENTION: Rat Array
: FILE REFERENCE: CLON-006CIP12
: CURRENT APPLICATION NUMBER: US/09/454,226A
: CURRENT FILING DATE: 2002-05-07
: NUMBER OF SEQ ID NOS: 1186
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 557
: LENGTH: 180
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Synthesized gene fragment
US-09-454-226A-557

Query Match 24.1% Score 88.8; DB 5; Length 180;
Best Local Similarity 68.3%; Pred. No. 2.1e-18;
Matches 123; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 101 tgttgagctctatctgttaacactgagactcctccctgggaagagagagccaaatgttaca 160
Db 1 tgcctcagtttagcaggttaacccctccagctcaggtgagtggaatggaataacgccttaattgccta 60
Qy 101 atgaacttaattggtgacccaagatatatgacctgtctgtgggactgtggaataactt 220

Db 160 TGTCCAGATGTCACCTGCTGCGGACTGATGGCTCACATATACCAATGAATGC 101
 Oy 235 gtgttatgtttgaaggtcggaacgacgacttctatctcattcaaaaaatctgggcct 294
 Db 100 CAGCTCTGCTTGGCCCGGATATAAAACCAACAGGACATCCAGATCATGAAGAATGGCAAA 41
 Oy 295 tgcctgagaaccaagg 309
 Db 40 TGCTGATCCACAGG 26

RESULT 8
 PCT-US01-43704-2330
 : Sequence 2330, Application PC/TUS0143704
 : GENERAL INFORMATION:
 : APPLICANT: Corixa Corporation
 : APPLICANT: Stolck, John A.
 : APPLICANT: Xu, Jiangchun
 : APPLICANT: Chenault, Ruth A.
 : APPLICANT: Meagher, Madelein Joy
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 : TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
 : FILE REFERENCE: 210121.561PC
 : CURRENT APPLICATION NUMBER: PCT/US01/43704
 : CURRENT FILING DATE: 2001-11-19
 : NUMBER OF SEQ ID NOS: 2606
 : SOFTWARE: Corixa Invention Disclosure Database
 : SEQ ID NO 2330
 : LENGTH: 323
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 PCT-US01-43704-2330

Query Match 11.5%; Score 42.2; DB 1; Length 323;
 Best Local Similarity 57.0%; Pred. No. 0.0019;
 Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 Oy 175 tgcacaaagatatgacctgtctgtgggactgtggaatacttctcattcaaaatctgggcct 234
 Db 164 tgtccagatgtccaaacctgctgctggcactgtggtctcacatatacgaatgaatgc 223
 Oy 235 gtgttatgtttgaaggtcggaacgacgacttctatctcattcaaaatctgggcct 294
 Db 224 cagctctgtgtggccggataaaacacacagacatccagatcatgaagaatggcaaa 283
 Oy 295 tgcctgagaaccaagg 309
 Db 284 tgctgatccacagg 298

RESULT 9
 US-09-671-483-2
 : Sequence 2, Application US/09671483
 : GENERAL INFORMATION:
 : APPLICANT: Kato, Seishi
 : APPLICANT: Yamaguchi, Tomoko
 : APPLICANT: Sekine, Shingo
 : APPLICANT: Kamata, Kouju
 : TITLE OF INVENTION: HUMAN PEC-60-LIKE PROTEIN AND DNA ENCODING THIS
 : TITLE OF INVENTION: PROTEIN
 : FILE REFERENCE: 6700PCT-US
 : CURRENT APPLICATION NUMBER: US/09/671.483
 : CURRENT FILING DATE: 2000-09-27
 : PRIOR APPLICATION NUMBER: US/09/065,019
 : PRIOR FILING DATE: 1998-04-17
 : NUMBER OF SEQ ID NOS: 3
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 2
 : LENGTH: 398
 : TYPE: DNA
 : ORGANISM: Homo sapiens

US-09-671-483-2
 Query Match 11.5%; Score 42.2; DB 5; Length 398;
 Best Local Similarity 57.0%; Pred. No. 0.002;
 Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 Oy 175 tgcacaaagatatgacctgtctgtgggactgtggaatacttctcattcaaaatctgggcct 234
 Db 178 tgtccagatgtccaaacctgctgctggcactgtggtctcacatatacgaatgaatgc 237
 Oy 235 gtgttatgtttgaaggtcggaacgacgacttctatctcattcaaaatctgggcct 294
 Db 238 cagctctgtgtggccggataaaacacacagacatccagatcatgaagaatggcaaa 297
 Oy 295 tgcctgagaaccaagg 309
 Db 298 tgctgatccacagg 312

RESULT 10
 US-10-106-698-1897
 : Sequence 1897, Application US/10106698
 : GENERAL INFORMATION:
 : APPLICANT: Ruben et al.
 : TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
 : FILE REFERENCE: P400SP1
 : CURRENT APPLICATION NUMBER: US/10/106,698
 : CURRENT FILING DATE: 2002-03-27
 : PRIOR APPLICATION NUMBER: PCT/US00/26524
 : PRIOR FILING DATE: 2000-09-28
 : PRIOR APPLICATION NUMBER: US 60/157,137
 : PRIOR FILING DATE: 1999-09-29
 : PRIOR APPLICATION NUMBER: US 60/163,280
 : PRIOR FILING DATE: 1999-11-03
 : NUMBER OF SEQ ID NOS: 8564
 : SOFTWARE: Patent In Ver. 3.0
 : SEQ ID NO 1897
 : LENGTH: 433
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-10-106-698-1897

Query Match 11.5%; Score 42.2; DB 7; Length 433;
 Best Local Similarity 57.0%; Pred. No. 0.0021;
 Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 Oy 175 tgcacaaagatatgacctgtctgtgggactgtggaatacttctcattcaaaatctgggcct 234
 Db 194 tgtccagatgtccaaacctgctgctggcactgtggtctcacatatacgaatgaatgc 253
 Oy 235 gtgttatgtttgaaggtcggaacgacgacttctatctcattcaaaatctgggcct 294
 Db 254 cagctctgtgtggccggataaaacacacagacatccagatcatgaagaatggcaaa 313
 Oy 295 tgcctgagaaccaagg 309
 Db 314 tgctgatccacagg 328

RESULT 11
 PCT-US02-10824-19
 : Sequence 19, Application PC/TUS0210824
 : GENERAL INFORMATION:
 : APPLICANT: Origene Technologies
 : TITLE OF INVENTION: Prostate Cancer Expression Profiles
 : FILE REFERENCE: 9U 206 PCT
 : CURRENT APPLICATION NUMBER: PCT/US02/10824
 : CURRENT FILING DATE: 2002-04-08
 : PRIOR APPLICATION NUMBER: US 60/281,732
 : PRIOR FILING DATE: 2001-04-06
 : PRIOR APPLICATION NUMBER: US 60/281,731

```
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 3528
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-10824-19

Query Match      11.0%; Score 40.6; DB 1; Length 3528;
Best Local Similarity 57.5%; Pred. No. 0.013;
Matches 73; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 170 atggatgcccaagaatataatgacctgtctgtggactgatggaaatacttaccctc 229
    || || || || || || || || || || || || || || || || || || || ||
Db 3054 atcttgtccaaggatttaaaagcctgtctgtggactgatggaaatacttaccctc 3113
    || || || || || || || || || || || || || || || || || || || ||

QY 230 aatgcgtgtattgttgaagtcggaacgccagacttctctcattcaaaaatctg 289
    || || || || || || || || || || || || || || || || || || || ||
Db 3114 ctgcatgctctgtcatgaaacccgtatagcccaaaaatacacacatccgcagtacag 3173
    || || || || || || || || || || || || || || || || || || || ||

QY 290 ggccttg 296
    || ||
Db 3174 ggaagtg 3180

RESULT 12
US-09-926-820-7
; Sequence 7, Application US/09926820
; GENERAL INFORMATION:
; APPLICANT: FORSSMANN, WOLF-GEORG
; APPLICANT: MAGERT, HANS-JURGEN
; APPLICANT: STANDKER, LUDGER
; APPLICANT: KREUTZMANN, PETER
; TITLE OF INVENTION: SERINE PROTEASE INHIBITORS
; FILE REFERENCE: 10496-P67431US0
; CURRENT APPLICATION NUMBER: US/09/926,820
; PRIOR FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: PCT/EP99/04331
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 3532
; TYPE: DNA
; ORGANISM: mammalian
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (44)...(3235)
US-09-926-820-7

Query Match      11.0%; Score 40.6; DB 5; Length 3532;
Best Local Similarity 57.5%; Pred. No. 0.013;
Matches 73; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 170 atggatgcccaagaatataatgacctgtctgtggactgatggaaatacttaccctc 229
    || || || || || || || || || || || || || || || || || || || ||
Db 3054 atcttgtccaaggatttaaaagcctgtctgtggactgatggaaatacttaccctc 3113
    || || || || || || || || || || || || || || || || || || || ||

QY 230 aatgcgtgtattgttgaagtcggaacgccagacttctctcattcaaaaatctg 289
    || || || || || || || || || || || || || || || || || || || ||
Db 3114 ctgcatgctctgtcatgaaacccgtatagcccaaaaatacacacatccgcagtacag 3173
    || || || || || || || || || || || || || || || || || || || ||

QY 290 ggccttg 296
    || ||
Db 3174 ggaagtg 3180

RESULT 13
US-09-926-820-12
; Sequence 12, Application US/09926820
; GENERAL INFORMATION:
; APPLICANT: FORSSMANN, WOLF-GEORG
; APPLICANT: MAGERT, HANS-JURGEN
; APPLICANT: STANDKER, LUDGER
; APPLICANT: KREUTZMANN, PETER
; TITLE OF INVENTION: SERINE PROTEASE INHIBITORS
; FILE REFERENCE: 10496-P67431US0
; CURRENT APPLICATION NUMBER: US/09/926,820
; PRIOR FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: PCT/EP99/04331
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 156
; TYPE: DNA
; ORGANISM: mammalian
US-09-926-820-12

Query Match      11.0%; Score 40.4; DB 5; Length 156;
Best Local Similarity 58.2%; Pred. No. 0.0056;
Matches 71; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 170 atggatgcccaagaatataatgacctgtctgtggactgatggaaatacttaccctc 229
    || || || || || || || || || || || || || || || || || || || ||
Db 32 atcttgtccaaggatttaaaagcctgtctgtggactgatggaaatacttaccctc 91
    || || || || || || || || || || || || || || || || || || || ||

QY 230 aatgcgtgtattgttgaagtcggaacgccagacttctctcattcaaaaatctg 289
    || || || || || || || || || || || || || || || || || || || ||
Db 52 ttgcatgctctgtcatgaaacccgtatagcccaaaaatacacacatccgcagtacag 151
    || || || || || || || || || || || || || || || || || || || ||

QY 290 qg 291
    ||
Db 152 qg 153

RESULT 14
US-09-629-469A-8198/c
; Sequence 8198, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORI
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8198
```

LENGTH: 562
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (484)..(484)
OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
FEATURE:
NAME/KEY: misc_feature
LOCATION: (505)..(505)
OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
FEATURE:
NAME/KEY: misc_feature
LOCATION: (552)..(552)
OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
FEATURE:
NAME/KEY: misc_feature
LOCATION: (560)..(560)
OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
FEATURE:
NAME/KEY: misc_feature
LOCATION: (562)..(562)
OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
US-09-629-469A-8198

Query Match 10.7% Score 39.2; DB 5; Length 562;
Best Local Similarity 58.6%; Pred. No. 0.021;
Matches 68; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
Qy 175 tgcaccaagatatatgaccctgtctgtgggactgatggaataacttaccatgaatgc 234
Db 335 TGCCCCAATGTGACAGCACCTGTTGTGCTCAAAATGGCCACACTTTCACAGAAATGAGTGT 275
Qy 235 gtgttatgtttgaaggctggaaacgcagacttctatctcctcaaaaaatctgg 290
Db 275 TTCCTTTGTGTGAACAGAGGGAATTCATTATCGTATAAAATTTGAAAAATATGG 220

RESULT 15
US-09-629-469A-11819
Sequence 11819, Application US/09629469A
GENERAL INFORMATION:
APPLICANT: OTA, TOSHIO
APPLICANT: ISOGAI, TAKAO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: HAYASHI, KOJI
APPLICANT: SAITO, KAORU
APPLICANT: YAMAMOTO, JUNICHI
APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OTSUKI, TETSUJI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
FILE REFERENCE: 084335/0123
CURRENT APPLICATION NUMBER: US/09/629,469A
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/159,590
PRIOR FILING DATE: 1995-10-18
PRIOR APPLICATION NUMBER: 60/183,322
PRIOR FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11819
LENGTH: 1022
TYPE: DNA
ORGANISM: Homo sapiens
US-09-629-469A-11819

Query Match 10.7% Score 39.2; DB 5; Length 1022;
Best Local Similarity 58.6%; Pred. No. 0.025;
Matches 68; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
Qy 175 tgcaccaagatatatgaccctgtctgtgggactgatggaataacttaccatgaatgc 234
Db 688 tgcccaaatgtgacagcacctgttggcctcaaatggccacacttccagaatgagtgt 747
Qy 235 gtgttatgtttgaaggctggaaacgcagacttctatctcctcaaaaaatctgg 290
Db 748 tctctttgtgtgaacagagggaatttcattatcgtataaaaaatttgaaaaatatgg 803

Search completed: July 29, 2002, 19:27:21
Job time: 7612 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 17:54:40 ; Search time 100.77 seconds
(without alignments)
897.024 Million cell updates/sec

Title: US-09-880-107-3847

Perfect score: 368
Sequence: 1 gaagagacgtgtaagtgcg.....aataaatgtatctgaatc 368

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NA.*

- 1: /cgn2.6/ptodata/2/ina/5A-COMB.seq.*
- 2: /cgn2.6/ptodata/2/ina/5B-COMB.seq.*
- 3: /cgn2.6/ptodata/2/ina/6A-COMB.seq.*
- 4: /cgn2.6/ptodata/2/ina/6B-COMB.seq.*
- 5: /cgn2.6/ptodata/2/ina/PTUS-COMB.seq.*
- 6: /cgn2.6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	211.4	57.4	265	US-09-221-298-55	Sequence 55, Appl
2	162.2	44.1	420	US-09-221-298-56	Sequence 56, Appl
3	162	44.0	302	US-08-235-515A-26	Sequence 26, Appl
4	109	29.6	458	US-08-880-829-21	Sequence 21, Appl
5	42.2	11.5	388	US-08-744-670-2	Sequence 2, Appl
6	42.2	11.5	388	US-09-149-933-2	Sequence 2, Appl
7	42.2	11.5	398	US-09-065-019-2	Sequence 2, Appl
8	38.2	10.4	180	US-09-065-019-1	Sequence 1, Appl
9	36	9.8	1565	US-08-468-846-1	Sequence 1, Appl
10	36	9.8	1565	US-08-915-096A-1	Sequence 1, Appl
11	35.4	9.6	126	US-08-586-676E-4	Sequence 4, Appl
12	35.4	9.6	129	US-08-586-676E-5	Sequence 5, Appl
13	35.4	9.6	138	US-08-586-676E-6	Sequence 6, Appl
14	34.6	9.4	149	US-08-586-676E-7	Sequence 7, Appl
15	33.8	9.2	149	US-08-586-676E-8	Sequence 8, Appl
16	33.4	9.1	4800	PCT-US94-07779-1	Sequence 1, Appl
17	33	9.0	1282	US-08-211-942-16	Sequence 16, Appl
18	32.8	8.9	309	US-08-839-709-2	Sequence 2, Appl
19	32.8	8.9	309	US-09-204-859-2	Sequence 2, Appl
20	31	8.4	2728	US-09-025-769B-299	Sequence 299, App
21	30.8	8.4	1541	US-08-446-924-1	Sequence 1, Appl
22	30.8	8.4	1541	US-08-798-665-1	Sequence 1, Appl
23	30.8	8.4	1541	US-08-982-987A-1	Sequence 1, Appl
24	30.6	8.3	1896	US-09-345-468-24	Sequence 24, Appl
25	30.6	8.3	2170	US-09-345-468-11	Sequence 11, Appl
26	30	8.2	3195	US-08-951-648-5	Sequence 5, Appl
27	30	8.2	3195	US-09-174-437-5	Sequence 5, Appl

28 29 7.9 908 4 US-09-439-313-350 Sequence 350, Appl
29 29 7.9 1122 2 US-08-820-170A-26 Sequence 26, Appl
30 29 7.9 1122 3 US-09-055-699-26 Sequence 26, Appl
31 29 7.9 1122 4 US-09-273-565-26 Sequence 26, Appl
32 29 7.9 1122 4 US-09-565-538-26 Sequence 26, Appl
33 29 7.9 1346 3 US-08-949-202-3 Sequence 3, Appl
34 29 7.9 1346 4 US-09-418-175-3 Sequence 3, Appl
35 29 7.9 1695 5 PCT-US95-06385-1 Sequence 1, Appl
36 29 7.9 1721 2 US-08-820-170A-27 Sequence 27, Appl
37 29 7.9 1721 3 US-09-055-699-27 Sequence 27, Appl
38 29 7.9 1721 4 US-09-273-565-27 Sequence 27, Appl
39 29 7.9 1721 4 US-09-565-538-27 Sequence 27, Appl
40 28.8 7.8 1057 3 US-09-188-930-18 Sequence 18, Appl
41 28.8 7.8 2004 3 US-09-188-930-230 Sequence 230, App
42 28.8 7.8 12839 3 US-09-125-287-1 Sequence 1, Appl
43 28.6 7.8 6765 2 US-08-677-010-2 Sequence 2, Appl
44 28.6 7.8 6765 2 US-08-790-519-2 Sequence 2, Appl
45 28.6 7.8 9581 2 US-08-677-010-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-221-298-55
: Sequence 55, Application US/09221298
: Patent No. 6284241
: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS
: TITLE OF INVENTION: OF COLON CANCER
: FILE REFERENCE: 210121.471
: CURRENT APPLICATION NUMBER: US/09/221,298
: CURRENT FILING DATE: 1998-12-23
: NUMBER OF SEQ ID NOS: 112
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 55
: LENGTH: 265
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (25)
: OTHER INFORMATION: Where n is a, c, g or t
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (223)
: OTHER INFORMATION: Where n is a, c, g or t
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (241)
: OTHER INFORMATION: Where n is a, c, g or t
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (254)
: OTHER INFORMATION: Where n is a, c, g or t
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (259)
: OTHER INFORMATION: Where n is a, c, g or t
US-09-221-298-55

Query Match 57.4%; Score 211.4; DB 4; Length 265;
Best Local Similarity 95.6%; Pred. NO. 4.2e-64;
Matches 237; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

Oy 76 atctttcttcagtccttgccctgttgatctatctgttaacacactggagctgactcc 135

Db 1 atctttcttcagtccttgccctgttgatctatctgttaacacactggagctgactcc 60

Oy 136 ctgggaagagagcgcacaaatgttacaatgaacttaagtatgacccaagatatatgacct 195

Tue Jul 30 07:47:47 2002

Db 61 ctggagagagagccaaatgttacaatgaacttaagtgaatgcacccaagatatgacccct 120
Qy 196 gtctgtggagactgatggaataactatcccaatgaatg-cgtgttatg-ttttgaaggtc 253
Db 121 gtctgtggagactgatggaataactatcccaatgaatg-cgtgttatg-ttttgaaggtc 180
Qy 254 gaaacgcagagact 313
Db 181 gaaacgcagagact 240
Qy 314 gaaatccc 321
Db 241 naaataccc 248

RESULT 2
US-09-221-298-56/c
; Sequence 56, Application US/09221298
; Patent No. 6284241
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471
; CURRENT APPLICATION NUMBER: US/09/221,298
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 56
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (85)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (164)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (198)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (228)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (257)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (261)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (283)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (298)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (326)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (385)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:

; NAME/KEY: modified_base
; LOCATION: (390)
; OTHER INFORMATION: Where n is a, c, g or t
US-09-221-298-56
Query Match 44.1%; Score 162.2; DB 4; Length 420;
Best Local Similarity 96.2%; Pred. No. 6.5e-47;
Matches 175; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Qy 160 aatgaacttaagtatgca-ccaagatatatgacctgtctgtggagactgatggaataac 218
Db 200 AANGAACITTAATGATGATGACCCCAAGATATATGACCCNGTCTGTGGGACTGATGGAATAAC 141
Qy 219 ttatcccaatgaatgcgtgttatgttttgaagtcggaacccagacactctctctctat 278
Db 140 TTATCCCAATGAATGCGTGTATGTTTGAATAATCGGAACGCCAGACTTCTATCTCAT 81
Qy 279 tcaaaaaatctggccttgcgtgagaaacaaaggttttgaataatcccatcaggtcacccgcgag 338
Db 80 TCAAAAATCTGGGCTTGGCTGCTGAGAACCAAGGTTTTTGAAATCCCATCAGGTACCCGCGAGG 21
Qy 339 cc 340
Db 20 AC 19
RESULT 3
US-08-235-315A-26
; Sequence 26, Application US/08235515A
; Patent No. 5840518
; GENERAL INFORMATION:
; APPLICANT: Morishita, Hideaki
; APPLICANT: Kanamori, Toshinori
; APPLICANT: No. 5840518uhara, Masahiro
; TITLE OF INVENTION: DNA FRAGMENT, VECTOR CONTAINING THE DNA
; TITLE OF INVENTION: FRAGMENT, TRANSFORMANT TRANSFORMED WITH THE VECTOR AND
; TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEIN USING THE VECTOR
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,515A
; FILING DATE: 29-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1110-135P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 1..302
; LOCATION: 1..302
; OTHER INFORMATION: /label= polynucleotide

OTHER INFORMATION: /note= "j-xl-y-z(psti), insert in plasmid pm474,
FIGURE INFORMATION: Figure 5
FEATURE: sig_peptide
NAME/KEY: sig_peptide
LOCATION: 27..89
FEATURE: mat_peptide
NAME/KEY: mat_peptide
LOCATION: 90..293
FEATURE: CDS
NAME/KEY: CDS
LOCATION: 27..293
US-08-235-515A-26

Query Match 44.0%; Score 162; DB 2; Length 302;
Best Local Similarity 91.9%; Pred. No. 6.5e-47;
Matches 171; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 120 cactgagctgactccctgggaagagcccaatgtttacaaacttaataatgaatgcac 179
Db 116 CTCGGGAATGGACTCCCTAGGTGGCGAGGCCAAATGTTACAAATGAACCTTAATGGATGCAC 175
Qy 180 caagatatatgacctgtctgtggagctgatggaatacttaccatgaatgcgtgtt 239
Db 176 CAAGATATATGACCTGCTGTGGGACTGATGGAATACTTATCCCAATGAATGCGTGT 235
Qy 240 atgttttgaagtggaaacgcagactctctatctcattcaaaaatctggcccttgcgtg 299
Db 236 ATGTTTGAATAATCGGAACGCCAGACATCGATCCTCATTTCAAAAATCTGGCCCTTGCTG 295
Qy 300 agaac 305
Db 296 AGGATC 301

RESULT 4
US-08-880-829-21
Sequence 21, Application US/08880829
Patent No. 5925559
GENERAL INFORMATION:
APPLICANT: Collins, John
TITLE OF INVENTION: A Collection of Phagemids, A
TITLE OF INVENTION: Collection of Escherichia Coli
TITLE OF INVENTION: Cells Carrying The Phagemids, A
TITLE OF INVENTION: Collection of Phagemid Particles
TITLE OF INVENTION: Produced From Said Collection
TITLE OF INVENTION: And Phagemid Particles
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Joseph T. Eisele
ADDRESSEE: Kane, Dalsimer, Sullivan, Kurucz,
ADDRESSEE: Levy, Eisele and Richard
STREET: 711 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10017-4059
COMPUTER READABLE FORM:
MEDIUM TYPE: 3-1/2" DISKETTE
COMPUTER: IBM-XT COMPATIBLE
OPERATING SYSTEM: DOS 3.3:
SOFTWARE: WORDPERFECT 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,829
FILING DATE: 23-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/458,668
FILING DATE: 06/02/95
APPLICATION NUMBER: German EP 94 108 689.4
FILING DATE: 06/07/94
ATTORNEY/AGENT INFORMATION:

NAME: EISELE, JOSEPH T.
REGISTRATION NUMBER: 25,331
REFERENCE/DOCKET NUMBER: 2727-77
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 687-6000
TELEFAX: (212) 682-3485
TELEX: (212) 426767
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 458 base pairs
TYPE: nucleic acid
STRANDEDNESS: single strand
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
US-08-850-329-21
Query Match 29.6%; Score 109; DB 2; Length 458;
Best Local Similarity 69.5%; Pred. No. 2.1e-28;
Matches 149; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
Qy 58 agtgccttgccctgtgagctatctgttaacactgagctgactccctgggaagagag 147
Db 133 ATTGCAGTGGCAGCTGGCTGTTTCCTACCGTAGCGAGCGCGACTCTCTGGTGGTAA 192
Qy 143 gccaaatgttacaatgaacttaattggtatgcaccaagatatatgacctgtctgtggagact 207
Db 133 GCTAAATGCTACACGAGCTCAACGGTTGCACTAAGATCTACGACCCGCTTTCGCGTACC 252
Qy 208 gatgaaatactattcccaatgaatgcgtgttatgttttgaagtcgaaacgccagact 267
Db 253 GACGGCAACACTTACCCGAAACGAATGCGTCTGTCTTCGAAACCCGTTAAACGTCAGACT 312
Qy 268 tctatcctcattcaaaaatctgggacctgtgcta 300
Db 313 TCTATCCTGATCCAGAAATCTGGTCCGTGCTTA 345

RESULT 5
US-08-744-570-2
Sequence 2, Application US/08744670
Patent No. 5858710
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: TUMOR-ASSOCIATED KAZAL INHIBITOR
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

```

: SOFTWARE: FastSeq Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/744,670
: FILING DATE: Filed Herewith
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0155 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 388 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: SINTTUT01
: CLONE: 1539065
: US-09-149-933-2

Query Match 11.5%; Score 42.2; DB:2; Length 388;
Best Local Similarity 57.0%; Pred. No. 3.1e-05;
Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Oy 175 tgcaccaagatatgacccctgtctgtgggactgtggaataacttaccatgaatgc 234
Db 168 TGTTCACAGATGTCACACTGCTCTGGCGCACTGATGGCTCACATATACGAATGAATGC 227
Oy 235 gtgttatgttttgaagtcggaacgccagacttctctctcattcaaaaatctgggcct 294
Db 228 CAGCTCTGCTTGGCCGGATATAAAACCAACAGGACATCCAGATCATGAAAGATGGCAAA 287
Oy 295 tctgagaaccaaagg 309
Db 288 TGCTGATCCACAGG 302

RESULT 7
US-09-065-019-2
: Sequence 2, Application US/09065019
: Patent No. 6280968
: GENERAL INFORMATION:
: APPLICANT: Kato, Seishi
: APPLICANT: Yamaguchi, Tomoko
: APPLICANT: Sekine, Shingo
: APPLICANT: Kamata, Kouju
: TITLE OF INVENTION: HUMAN PEC-60-LIKE PROTEIN AND DNA ENCODING THIS PROTEIN
: FILE REFERENCE: 6700PCT-US
: CURRENT APPLICATION NUMBER: US/09/065,019
: CURRENT FILING DATE: 1998-04-17
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 398
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-065-019-2

Query Match 11.5%; Score 42.2; DB 4; Length 398;
Best Local Similarity 57.0%; Pred. No. 3.2e-05;
Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Oy 175 tgcaccaagatatgacccctgtctgtgggactgtggaataacttaccatgaatgc 234
Db 178 tgttccacagatgtccaaacctggctcgtggcactgagtggtcgcacatatacgaatgaatgc 237
Oy 235 gtgttatgttttgaagtcggaacgccagacttctctctcattcaaaaatctgggcct 294
Db 238 cagctctgttggcccgataaaacacacagacacagacatccagatcatgaagaatggcaaa 297
Oy 295 tctgagaaccaaagg 309
Db 298 tgctgatcccaagg 312

```

```

: SOFTWARE: FastSeq Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/744,670
: FILING DATE: Filed Herewith
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0155 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 388 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: SINTTUT01
: CLONE: 1539065
: US-08-744-670-2

Query Match 11.5%; Score 42.2; DB 2; Length 388;
Best Local Similarity 57.0%; Pred. No. 3.1e-05;
Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Oy 175 tgcaccaagatatgacccctgtctgtgggactgtggaataacttaccatgaatgc 234
Db 168 TGTTCACAGATGTCACACTGCTCTGGCGCACTGATGGCTCACATATACGAATGAATGC 227
Oy 235 gtgttatgttttgaagtcggaacgccagacttctctctcattcaaaaatctgggcct 294
Db 228 CAGCTCTGCTTGGCCGGATATAAAACCAACAGGACATCCAGATCATGAAAGATGGCAAA 287
Oy 295 tctgagaaccaaagg 309
Db 288 TGCTGATCCACAGG 302

RESULT 6
US-09-149-933-2
: Sequence 2, Application US/09149933
: Patent No. 5958699
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Goli, Surya K.
: APPLICANT: Murry, Lynn E.
: TITLE OF INVENTION: TUMOR-ASSOCIATED KAZAL INHIBITOR
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: US
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/149,933
: FILING DATE: Filed Herewith
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.

```

```

RESULT      8
US-09-065-019-1
: Sequence 1, Application US/09065019
: Patent No. 6280988
: GENERAL INFORMATION:
: APPLICANT: Kato, Seishi
: APPLICANT: Yamaguchi, Tomoko
: APPLICANT: Sekine, Shingo
: APPLICANT: Kamata, Kouju
: TITLE OF INVENTION: HUMAN PEC-60-LIKE PROTEIN AND DNA ENCODING THIS PROTEIN
: FILE REFERENCE: 6700PCT-US
: CURRENT APPLICATION NUMBER: US/09/065,019
: CURRENT FILING DATE: 1998-04-17
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 1
: LENGTH: 180
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-065-019-1

```

	Query Match	10.43;	Score 38.2;	DB 4;	Length 180;
Best Local	Similarity	56.94;	Pred. No. 0.00054;		
Matches	70;	Conservative	0;	Mismatches 53;	Indels 0;
Qy 175	tgaccaagatata	gacctgtctgtgggactgatggaatactattcccaatgaatgc	234		
Db 58	tgtcccagatgcc	aacctgtctgtggcactgatggctcacatcacgaatgaatgc	117		
Qy 235	gtgttatgttttg	aaggtcggaaacgcgcagactctctatccctattcaaaaatctgggcct	294		
Db 118	caagtctcttcgc	cccgatcaaaaacacacagacatccagatcatgaagatgcgcaaa	177		

Qy	295	tgc	297
Dd	178	tgc	180

```

1  RESULT          9
2  US-08-468-846-1
3  : Sequence 1, Application US/08468846
4  : Patent No. 6074839
5  : GENERAL INFORMATION:
6  : APPLICANT: Meissner, Paul
7  : APPLICANT: Fuldner, Rebecca
8  : APPLICANT: Fei-wei, Ying
9  : APPLICANT: Adams, Mark
10 : TITLE OF INVENTION: TRANSFORMING GROWTH FACTOR ALPHA HI
11 :
12 : NUMBER OF SEQUENCES: 15
13 : CORRESPONDENCE ADDRESSES:
14 : ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
15 : ADDRESSEE: STUART & OLSTEIN
16 : STREET: 6 Becker Farm Road
17 : CITY: Roseland
18 : STATE: New Jersey
19 : COUNTRY: USA
20 : ZIP: 07068
21 :
22 : COMPUTER READABLE FORM:
23 : MEDIUM TYPE: Floppy disk
24 : COMPUTER: IBM PC compatible
25 : OPERATING SYSTEM: PC-DOS/MS-DOS
26 : SOFTWARE: PatentIn Release #1.0, Version #1.30
27 : CURRENT APPLICATION DATA:
28 : APPLICATION NUMBER: US/08/468,846
29 : FILING DATE: 06-JUN-1995
30 : CLASSIFICATION: 514
31 :
32 : PRIOR APPLICATION DATA:
33 : APPLICATION NUMBER: US 08/208,008
34 : FILING DATE: 08-MAR-1994
35 : ATTORNEY/AGENT INFORMATION:
36 : NAME: Ferraro, Gregory D.
37 : REGISTRATION NUMBER: 36,134

```

```

: REFERENCE/DOCKET NUMBER: 325800-465
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-994-1700
: TELEFAX: 201-994-1744
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1565 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 68..1207
: US-08-468-846-1

Query Match          9.8%; Score 36; DB 3; Length 1565;
Best Local Similarity 55.8%; Pred. No. 0.0085;
Matches 69; Conservative

Qy 175 tgaccaagatataatgacctgtctgtggagactgtggaataactatcccaatgaatgc 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 TGCCATACAAATTATATTCTGTCTGTGGATCAAAATGGGACACTTATCAAAATGAATGC 433

Qy 235 gtgtatgttttgaaggtcgaaacgccagacttctatcctcattcaaaaatctggcct 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 434 TTCTCAGAAGGGCTGCTTGTGAAGCACCAGAAAGAGATACACTATATACGACAGGACCA 493

Qy 295 tgcct 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 494 TGCT 497

```

RESULT 10
US-08-915-096A-1
: Sequence 1, Application US/08915096A
: Patent No. 6265543
: GENERAL INFORMATION:
: APPLICANT: Weissner, Paul S.
: APPLICANT: Fuldner, Rebecca A.
: APPLICANT: Adams, Mark D.
: TITLE OF INVENTION: Transforming Growth Factor Alpha H1
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: MD
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/915,096A
: FILING DATE: 20-AUG-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/468,846
: FILING DATE: 06-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/208,008
: FILING DATE: 08-MAR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET INFORMATION: PF110D1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 301-309-8504
: TELEFAX: 301-309-8439

Query Match 9.6%; Score 35.4; DB 2; Length 126;
Best Local Similarity 38.3%; Pred. No. 0.0043;
Matches 41; Conservative 20; Mismatches 46; Indels

Query Match 9.6%; Score 35.4; DB 2; Length 126;
Best Local Similarity 38.3%; Pred. NO. 0.0043;
Matches 41; Conservative 20; Mismatches 46; Indels 0; Gaps

Qy	160	aatgaaccttaatggatgcaccagaatatatgacccctgtctgtggagactaatggaaatact	219
Db	1	AARAARGINTGYCNGTGYCCNAARATHYTNAAKCCNGTNTGYGNNWSNAGAYGNNMGNACN	60
Qy	220	tatcccaatgaatgcgtgttatgtttgaaggtcggaaacgcagac	266
Db	61	TAYGCNAAAYNSNTGYATHGCMGNTGYAAYGGNGTNKSNATHAARS	107

```

RESULT 12
US-08-586-576E-5
: Sequence 5, Application US/08586676E
: Patent No. 5972698
: GENERAL INFORMATION:
: APPLICANT: Fritz, Hans,
: APPLICANT: Sommerhoff, Christian
: TITLE OF INVENTION: Tryptase inhibitor
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 5972698Bartis Corporation, Patent and Trademark Department
: STREET: 564 Morris Avenue
: CITY: Summit
: STATE: New Jersey
: COUNTRY: US
: ZIP: 07901-1027
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/586.676E
: FILING DATE: 25-JAN-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/EP94/02445
: FILING DATE: 25-JUL-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: EP 9311930.9
: FILING DATE: 26-JUL-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Pfeiffer, Hessa J.
: REGISTRATION NUMBER: 22.640
: REFERENCE/DOCKET NUMBER: 4-20076/PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 522 6940
: TELEFAX: (908) 522 6955
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 129 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: YES
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Hirudo medicinalis
: US-08-586-676E-5

```

Query: match

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 19:20:56 ; Search time 4748.22 Seconds
(without alignments)
1676.764 Million cell updates/sec

Title: US-09-880-107-3847
Perfect score: 368
Sequence: 1 gaagaacgtgtaagtcg.....aataaatgtatctgaatc 368

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues 43959072
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main.*

- 1: /cgn2_6/ptodata/2/pna/US06005_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06006_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US080_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
- 9: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
- 10: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
- 11: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
- 12: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
- 13: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
- 14: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
- 15: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
- 16: /cgn2_6/ptodata/2/pna/US092_COMB.seq.*
- 17: /cgn2_6/ptodata/2/pna/US093_COMB.seq.*
- 18: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
- 19: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
- 20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
- 21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
- 22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
- 23: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
- 24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
- 25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
- 26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
- 27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
- 28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
- 29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
- 30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
- 31: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
- 32: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
- 33: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
- 34: /cgn2_6/ptodata/2/pna/US098D_COMB.seq.*
- 35: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
- 36: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
- 37: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
- 38: /cgn2_6/ptodata/2/pna/US100_COMB.seq.*
- 39: /cgn2_6/ptodata/2/pna/US101_COMB.seq.*
- 40: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
- 41: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
- 42: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
- 43: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*

- 44: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
- 45: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
- 46: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
- 47: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
- 48: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
- 49: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
- 50: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
- 51: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
- 52: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
- 53: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
- 54: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
- 55: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
- 56: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
- 57: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
- 58: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
- 59: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
- 60: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
- 61: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
- 62: /cgn2_6/ptodata/2/pna/US6023_COMB.seq.*
- 63: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
- 64: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
- 65: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
- 66: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
- 67: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
- 68: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
- 69: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
- 70: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
- 71: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
- 72: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
- 73: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
- 74: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
- 75: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	368	100.0	368	33	US-09-880-107-3847	Sequence 3847, Ap
2	368	100.0	368	35	US-09-954-456-91	Sequence 91, Appl
3	368	100.0	368	35	US-09-954-456-730	Sequence 730, Appl
4	368	100.0	368	68	US-60-298-951-5	Sequence 5, Appl
5	344.8	93.7	378	35	US-09-948-941-198	Sequence 198, Appl
6	344.8	93.7	396	16	US-09-248-797-17742	Sequence 17742, A
7	344.8	93.7	396	17	US-09-346-956-7285	Sequence 7285, Ap
8	344.8	93.7	396	34	US-09-904-703-7285	Sequence 7285, Ap
9	344.8	93.7	396	34	US-09-925-564-17742	Sequence 17742, A
10	344.8	93.7	604	30	US-09-925-564-17742	Sequence 16, Appl
11	344.8	93.7	604	1	PCT-US00-05989-292	Sequence 292, Appl
12	344.8	93.7	604	34	US-09-925-297-292	Sequence 292, Appl
13	344.8	93.7	607	30	US-09-760-477-217	Sequence 217, Appl
14	344.8	93.7	607	30	US-09-760-492-57	Sequence 57, Appl
15	344.8	93.7	608	25	US-09-647-801-11	Sequence 11, Appl
16	344.8	93.7	839	56	US-06-172-373-7243	Sequence 7243, Ap
17	344.8	93.7	937	25	US-09-652-127-9525	Sequence 9525, Ap
18	343.8	93.4	459	55	US-06-164-285-4397	Sequence 4397, Ap
19	342.6	93.1	591	25	US-09-652-127-7348	Sequence 7348, Ap
20	342.6	93.1	591	25	US-09-652-914-7443	Sequence 7443, Ap
21	340.8	92.6	432	14	US-09-016-434-1250	Sequence 1250, Ap
22	337.8	91.8	478	20	US-09-534-840-9544	Sequence 9544, Ap
23	334.8	91.0	348	75	US-06-363-352-2205	Sequence 2205, Ap
24	328	89.1	1115	33	US-09-881-797-3501	Sequence 3501, Ap
25	328	89.1	1115	37	US-10-021-698-3501	Sequence 3501, Ap
26	324.8	88.3	438	1	PCT-US01-08631-27303	Sequence 27303, A
27	322.6	87.7	379	1	PCT-US01-30732-1358	Sequence 1358, Ap
28	321.8	87.4	421	31	US-09-823-327-5264	Sequence 5264, Ap
29	316.6	86.0	362	1	PCT-US01-30732-4270	Sequence 4270, Ap
30	298	81.0	350	17	US-09-362-510-13590	Sequence 13590, A
31	298	81.0	350	17	US-09-362-510A-13590	Sequence 13590, A

32 298 81.0 350 34 US-09-904-013-13590 Sequence 13590, A
33 286.8 77.9 250 20 US-09-534-840-9542 Sequence 9542, Ap
34 280.6 76.3 378 31 US-09-824-130-2835 Sequence 2835, Ap
35 272.8 74.1 276 20 US-09-534-840-9543 Sequence 9543, Ap
36 266.2 72.3 271 17 US-09-304-517A-54874 Sequence 54874, A
37 266.2 72.3 271 17 US-09-371-146A-54874 Sequence 54874, A
38 266.2 72.3 271 36 US-09-985-678-54874 Sequence 54874, A
39 265 72.0 347 1 PCT-US01-30732-1734 Sequence 1734, Ap
40 264.6 71.9 369 31 US-09-823-327-3238 Sequence 3238, Ap
41 264.2 71.8 300 7 US-08-303-241-1785 Sequence 1785, Ap
42 264.2 71.8 300 20 US-09-534-840-10001 Sequence 10001, A
43 231.6 62.9 448 30 US-09-760-477-175 Sequence 125, App
44 223 60.6 510 74 US-60-350-061-74 Sequence 74, Appl
45 222.4 60.4 549 30 US-09-770-171-1053 Sequence 1053, Ap

ALIGNMENTS

RESULT .1
US-09-880-107-3847
: Sequence 3847, Application US/09880107
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880.107
: PRIOR FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 3847
: LENGTH: 368
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. Y00705
US-09-880-107-3847

Query Match 100.0%; Score 368; DB 33; Length 368;
Best Local Similarity 100.0%; Pred. No. 5.1e-106;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gaagagacgtggttaagtgcggtgcagttttcaactgacctctggacgcagaaacttcagcc 60
Db 1 gaagagacgtggttaagtgcggtgcagttttcaactgacctctggacgcagaaacttcagcc 60
QY 61 atgaaggttaacaggacatctttcttcagtcgcttgccctgttgatctatctggttaac 120
Db 61 atgaaggttaacaggacatctttcttcagtcgcttgccctgttgatctatctggttaac 120
QY 121 actggagctgactccctgggaagagagggccaaatgttacaatgaacttaattggtgcacc 180
Db 121 actggagctgactccctgggaagagagggccaaatgttacaatgaacttaattggtgcacc 180
QY 181 aagatatagacctgtctgtggagctgatgaaataacttaccctaaatgaatgcgtgtta 240
Db 181 aagatatagacctgtctgtggagctgatgaaataacttaccctaaatgaatgcgtgtta 240
QY 241 tgttttgaaggtcggaaacgccagactctctatctcatcaaaaaatctgggcttgcgtga 300
Db 241 tgttttgaaggtcggaaacgccagactctctatctcatcaaaaaatctgggcttgcgtga 300
QY 301 gaaccaaggttttgaataccatccaggtcacccgcgagggcctattgttgaataaatgtatc 360
Db 301 gaaccaaggttttgaataccatccaggtcacccgcgagggcctattgttgaataaatgtatc 360

QY 361 tgaataatc 368
Db 361 tgaataatc 368
RESULT .2
US-09-954-456-91
: Sequence 91, Application US/09954456
: GENERAL INFORMATION:
: APPLICANT: Young, Paul
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-76
: CURRENT APPLICATION NUMBER: US/09/954.456
: PRIOR FILING DATE: 2001-09-18
: PRIOR APPLICATION NUMBER: US/60/233,617
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,052
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,923
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,134
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,637
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,638
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,711
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,720
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,840
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,863
: PRIOR FILING DATE: 2000-09-27
: NUMBER OF SEQ ID NOS: 2276
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 91
: LENGTH: 368
: TYPE: DNA
: ORGANISM: Homo sapiens
: OTHER INFORMATION: US-09-954-456-91

Query Match 100.0%; Score 368; DB 35; Length 368;
Best Local Similarity 100.0%; Pred. No. 5.1e-106;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gaagagacgtggttaagtgcggtgcagttttcaactgacctctggacgcagaaacttcagcc 60
Db 1 gaagagacgtggttaagtgcggtgcagttttcaactgacctctggacgcagaaacttcagcc 60
QY 61 atgaaggttaacaggacatctttcttcagtcgcttgccctgttgatctatctggttaac 120
Db 61 atgaaggttaacaggacatctttcttcagtcgcttgccctgttgatctatctggttaac 120
QY 121 actggagctgactccctgggaagagagggccaaatgttacaatgaacttaattggtgcacc 180
Db 121 actggagctgactccctgggaagagagggccaaatgttacaatgaacttaattggtgcacc 180
QY 161 aagatatagacctgtctgtggagctgatgaaataacttaccctaaatgaatgcgtgtta 240
Db 181 aagatatagacctgtctgtggagctgatgaaataacttaccctaaatgaatgcgtgtta 240
QY 241 tgttttgaaggtcggaaacgccagactctctatctcatcaaaaaatctgggcttgcgtga 300
Db 241 tgttttgaaggtcggaaacgccagactctctatctcatcaaaaaatctgggcttgcgtga 300
QY 301 gaaccaaggttttgaataccatccatccaggtcacccgcgagggcctattgttgaataaatgtatc 360
Db 301 gaaccaaggttttgaataccatccatccaggtcacccgcgagggcctattgttgaataaatgtatc 360


```
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-198

Query Match      93.7%; Score 344.8; DB 35; Length 378;
Best Local Similarity 96.8%; Pred. No. 1.2e-98;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Qy 1 gaagagacgtggttaagtcggtgagctgtttcaactgaacacctctggagcagaaacttcagcc 60
   |||||||
Db 1 gaagagacgtggttaagtcggtgagctgtttcaactgaacacctctggagcagaaacttcagcc 60

Qy 61 atgaagtaacagcagctcttcttcagtcgcttggccttggcctgttgatctatctggttaac 120
   |||||||
Db 61 atgaagtaacagcagctcttcttcagtcgcttggccttggcctgttgatctatctggttaac 120

Qy 121 actggagctgactccctgggaagagagccaaatgttacaatgaacttaattgagtgacc 180
   |||||||
Db 121 actggagctgactccctgggaagagagccaaatgttacaatgaacttaattgagtgacc 180

Qy 181 aagatatgacctgtctgtgagctgagtgaaatacttaccatgaatgcgtgtta 240
   |||||||
Db 181 aagatatgacctgtctgtgagctgagtgaaatacttaccatgaatgcgtgtta 240

Qy 241 tgtttgaagtcggaacccagacacttctatctcattcaaaaaatctgggacctgtga 300
   |||||||
Db 241 tgtttgaagtcggaacccagacacttctatctcattcaaaaaatctgggacctgtga 300

Qy 301 gaaccaaggtttgaaatcccatcaggtccaccgagggcc-----tattgttga 350
   |||||||
Db 301 gaaccaaggtttgaaatcccatcaggtccaccgagggcc-----tattgttga 350

Qy 351 taaatgtatctgaatc 368
   |||||||
Db 351 taaatgtatctgaatc 368

Qy 361 taaatgtatctgaatc 378
   |||||||
Db 361 taaatgtatctgaatc 378

RESULT 6
US-09-248-797-17742
; Sequence 17742, Application US/09248797
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-764
; CURRENT APPLICATION NUMBER: US/09/248,797
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 48909
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17742
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-248-797-17742

Query Match      93.7%; Score 344.8; DB 16; Length 396;
Best Local Similarity 96.8%; Pred. No. 1.3e-98;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Qy 1 gaagagacgtggttaagtcggtgagctgtttcaactgaacacctctggagcagaaacttcagcc 60
   |||||||
Db 17 gaagagacgtggttaagtcggtgagctgtttcaactgaacacctctggagcagaaacttcagcc 76

Qy 61 atgaagtaacagcagctcttcttcagtcgcttggccttggcctgttgatctatctggttaac 120
   |||||||
Db 77 atgaagtaacagcagctcttcttcagtcgcttggccttggcctgttgatctatctggttaac 136

Qy 121 actggagctgactccctgggaagagagccaaatgttacaatgaacttaattgagtgacc 180
   |||||||
Db 137 actggagctgactccctgggaagagagccaaatgttacaatgaacttaattgagtgacc 196

Qy 181 aagatatgacctgtctgtgagctgagtgaaatacttaccatgaatgcgtgtta 240
   |||||||
Db 197 aagatatgacctgtctgtgagctgagtgaaatacttaccatgaatgcgtgtta 256

Qy 241 tgtttgaagtcggaacccagacacttctatctcattcaaaaaatctgggacctgtga 300
   |||||||
Db 257 tgtttgaagtcggaacccagacacttctatctcattcaaaaaatctgggacctgtga 316

Qy 301 gaaccaaggtttgaaatcccatcaggtccaccgagggcc-----tattgttga 350
   |||||||
Db 317 gaaccaaggtttgaaatcccatcaggtccaccgagggcc-----tattgttga 376

Qy 351 taaatgtatctgaatc 368
   |||||||
Db 377 taaatgtatctgaatc 394

RESULT 8
US-09-924-703-7285
```

Sequence 7285, Application US/09904703
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-758CON1
CURRENT APPLICATION NUMBER: US/09/904,703
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: 09/210,298
PRIOR FILING DATE: 1998-12-09
NUMBER OF SEQ ID NOS: 17812
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7285
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapiens
US-09-904-703-7285

Query Match 93.7%; Score 344.8; DB 34; Length 396;
Best Local Similarity 96.8%; Pred. No. 1.3e-98;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 1 gaagagacgtgtaagtcggtgcagttttcaactgacctctgacgcagaaacttcagcc 60
DB 17 gaagagacgtgtaagtcggtgcagttttcaactgacctctgacgcagaaacttcagcc 76
QY 61 atgaaggttaacaggcatctttctcagtcgcttggccctgttgagtcctatctggttaac 120
DB 77 atgaaggttaacaggcatctttctcagtcgcttggccctgttgagtcctatctggttaac 136
QY 121 actggagctgactccctgggaagagagcgaatattacaatgaacttaagtgaatgcacc 180
DB 137 actggagctgactccctgggaagagagcgaatattacaatgaacttaagtgaatgcacc 196
QY 181 aagatatgaacctctgtctgggactgatggaataacttatcccaatgaatgcgtgta 240
DB 197 aagatatgaacctctgtctgggactgatggaataacttatcccaatgaatgcgtgta 256
QY 241 tgtttgaaggtcggaacgacgaccttctctcattcattcaaaaatctgggcttgcgtga 300
DB 257 tgtttgaaggtcggaacgacgaccttctctcattcattcaaaaatctgggcttgcgtga 316
QY 301 gaaccaaggtttgaaatcccatcaggtcaccgcgagggcc-----tattgtgaa 350
DB 317 gaaccaaggtttgaaatcccatcaggtcaccgcgagggcc-----tattgtgaa 376
QY 351 taaatgtatctgaatc 368
DB 377 taaatgtatctgaatc 394

RESULT 10
US-09-750-492-16
Sequence 16, Application US/09760492
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PTZ67
CURRENT APPLICATION NUMBER: US/09/760,492
CURRENT FILING DATE: 2001-01-16
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 184
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 603
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIIE
LOCATION: (584)
OTHER INFORMATION: n equals a,t,g, or c
US-09-760-492-16

Query Match 93.7%; Score 344.8; DB 30; Length 603;
Best Local Similarity 96.8%; Pred. No. 1.5e-98;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 1 gaagagacgtgtaagtcggtgcagttttcaactgacctctgacgcagaaacttcagcc 60
DB 177 gaagagacgtgtaagtcggtgcagttttcaactgacctctgacgcagaaacttcagcc 236
QY 61 atgaaggttaacaggcatctttctcagtcgcttggccctgttgagtcctatctggttaac 120
DB 237 atgaaggttaacaggcatctttctcagtcgcttggccctgttgagtcctatctggttaac 296
QY 121 actggagctgactccctgggaagagagcgaatattacaatgaacttaagtgaatgcacc 180
DB 257 actggagctgactccctgggaagagagcgaatattacaatgaacttaagtgaatgcacc 356
QY 181 aagatatgaacctctgtctgggactgatggaataacttatcccaatgaatgcgtgta 240

Sequence 7285, Application US/09904703
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-758CON1
CURRENT APPLICATION NUMBER: US/09/904,703
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: 09/210,298
PRIOR FILING DATE: 1998-12-09
NUMBER OF SEQ ID NOS: 17812
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7285
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapiens
US-09-904-703-7285

Query Match 93.7%; Score 344.8; DB 34; Length 396;
Best Local Similarity 96.8%; Pred. No. 1.3e-98;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 1 gaagagacgtgtaagtcggtgcagttttcaactgacctctgacgcagaaacttcagcc 60
DB 17 gaagagacgtgtaagtcggtgcagttttcaactgacctctgacgcagaaacttcagcc 76
QY 61 atgaaggttaacaggcatctttctcagtcgcttggccctgttgagtcctatctggttaac 120
DB 77 atgaaggttaacaggcatctttctcagtcgcttggccctgttgagtcctatctggttaac 136
QY 121 actggagctgactccctgggaagagagcgaatattacaatgaacttaagtgaatgcacc 180
DB 137 actggagctgactccctgggaagagagcgaatattacaatgaacttaagtgaatgcacc 196
QY 181 aagatatgaacctctgtctgggactgatggaataacttatcccaatgaatgcgtgta 240
DB 197 aagatatgaacctctgtctgggactgatggaataacttatcccaatgaatgcgtgta 256
QY 241 tgtttgaaggtcggaacgacgaccttctctcattcattcaaaaatctgggcttgcgtga 300
DB 257 tgtttgaaggtcggaacgacgaccttctctcattcattcaaaaatctgggcttgcgtga 316
QY 301 gaaccaaggtttgaaatcccatcaggtcaccgcgagggcc-----tattgtgaa 350
DB 317 gaaccaaggtttgaaatcccatcaggtcaccgcgagggcc-----tattgtgaa 376
QY 351 taaatgtatctgaatc 368
DB 377 taaatgtatctgaatc 394

RESULT 9
US-09-925-564-17742
Sequence 17742, Application US/09925564
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-764
CURRENT APPLICATION NUMBER: US/09/925,564
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/248,797
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 48909
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17742
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-564-17742

Db 357 aagatatagacctgtctgtggactgatggaaatactatcccaatgaatcgtctta 415
QY 241 tgttttgaagtcggaacgcgcagacttctatctctatcaaaaatctgggccttgcga 300
Db 417 tgttttgaatacggaaacgccagacttctatctctatcaaaaatctgggccttgcga 476
QY 301 gaaccaagttttgaaatcccatccatcaggctcaccggaagcc-----tatgttgaa 350
Db 477 gaaccaagttttgaaatcccatccatcaggctcaccggaagcctgactggccttatgttgaa 536
QY 351 taaatgtatctgaatc 368
Db 537 taaatgtatctgaatc 554

```

RESULT 11
PCT-US00-05989-292
: Sequence 292, Application PC/TUS0005989
: GENERAL INFORMATION:
: APPLICANT: Craig Rosen,
: APPLICANT: Steve Ruben
: TITLE OF INVENTION: Human Pancreas and Pancreatic Cancer Associated Gene Sequences and
: FILE REFERENCE: PA105PCT
: CURRENT APPLICATION NUMBER: PCT/US00/05989
: CURRENT FILING DATE: 2000-03-08
: EARLIER APPLICATION NUMBER: 60/124,270
: EARLIER FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 928
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 292
: LENGTH: 604
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (557)
: OTHER INFORMATION: n equals a,t,g, or c
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (580)
: OTHER INFORMATION: n equals a,t,g, or c
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (582)
: OTHER INFORMATION: n equals a,t,g, or c
PCT-US00-05989-292

```

	Query Match	93.7%;	Score 344.8;	DB 1:	Length 604;
	Best Local Similarity	96.8%;	Pred. No. 1.5e-98;		
	Matches 366; Conservative	0;	Mismatches 2;	Indels	Gaps
Oy	1 gaagagacgttgtaagaatgcggtgcagtctttcaactgaacctctggacgcagaacttcagcc 60 				
Db	179 gaagagcgttgtaagaatgcggtgcagtctttcaactgaacctctggacgcagaacttcagcc 238 				
Oy	61 atgaaggataacaggcatcttcttcagtgcccttgacccttgtaacttatctcgtaaac 120 				
Db	239 atgaaggataacaggcatcttcttcagtgcccttgacccttgtaacttatctcgtaaac 298 				
Oy	121 actggagctgactccctgggaagagagccc aaatggttacaatgaacttaatggatgcc 180 				
Db	299 actggagctgactccctlgggaagagagccc aaatggttacaatgaacttaatggatgcc 358 				
Oy	181 aagatatagacctgctgtggagactgatgaaatacttatccccaatgaatcggttta 240 				
Db	359 aagatatagacctgctgtggagactgatgaaatacttatccccaatgaatcggttta 418 				
Oy	241 tgtttgaagatcggaacgccagactctctatccattcaaaaactcgggccttgtcta 300 				
Db	419 tottttaaagatcggaacgccagactctctatccattcaaaaactcgggccttgtcta 478 				

Qy	301	gaaccaagggtttgaaatcccatcaggtcacccgaggcc-----tattattgaa	350
Db	479	gaaccaagggtttgaaatcccatcaggtcacccgaggccattattgtgaa	538
Qy	351	taaagtatctgaataatc	368
Db	539	taaagtatctgaataatc	556

```

RESULT 12
US-09-925-297-292
: Sequence 292, Application US/09925297
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA105
: CURRENT APPLICATION NUMBER: US/09/925,297
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05989
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 928
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 292
: LENGTH: 604
: TYPE: -NA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (557)
: OTHER INFORMATION: n equals a.t.g. or c
: NAME/KEY: misc feature
: LOCATION: (580)
: OTHER INFORMATION: n equals a.t.g. or c
: NAME/KEY: misc feature
: LOCATION: (582)
: OTHER INFORMATION: n equals a.t.g. or c
US-09-925-297-292

```

Query March 93.7%; Score 344.8; DB 34; Length 604;
Best Local Similarity 96.8%; Pred. NO. 1.5e-98;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

Qy	1	gaagagacgttggttaagtgcggtgcaggttttcaactgacctctctggagcgcgagaacttcagcc	60
Db	179	gaagagacgttggttaagtgcggtgcaggttttcaactgacctctctggagcgcgagaacttcagcc	238
Qy	61	atgaaggttaacagcagcatcttctctcaagtccttgccctgttgagctctatctgttaac	120
Db	239	atgaaggttaacagcagcatcttctctcaagtccttgccctgttgagctctatctgttaac	298
Qy	121	actggagctgactccctgggaagagagccaaatgttacaatgaacttaatggatgcc	180
Db	299	actggagctgactccctgggaagagagccaaatgttacaatgaacttaatggatgcc	358
Qy	181	aagatatagacctgtctgttggaactgatggaataacttatcccaatgaatgcggtta	240
Db	359	aagatatagacctgtctgttggaactgatggaataacttatcccaatgaatgcggtta	418
Qy	241	tgttttggaaggtcggaaacgcgaacttctatccctatcaaaaactcgggccttgcga	3000
Db	419	tgttttgaaatcggaaacgcgcagacttctatccctatcaaaaactcgggccttgcga	478
Qy	301	gaaccaaagttttgaaatcccatcaggttcacgcgagggc-----tattgttaa	350
Db	479	gaaccaaggttttgaaatcccatcaggttcacgcgagggcctgactggccctattgttgaa	538
Qy	351	taaatgtatctgaatatc	368
Db	539	taaatgtatctgaatatc	556

RESULT 13

US-09-760-477-217/c
: Sequence 217, Application US/09760477
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC013
: CURRENT APPLICATION NUMBER: US/09760,477
: CURRENT FILING DATE: 2001-01-16
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 734
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 217
: LENGTH: 607
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (23)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (25)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (35)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-760-477-217

Query Match 93.7%; Score 344.8; DB 30; Length 607;

Best Local Similarity 96.8%; Pred. No. 1.5e-98;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 1 gaagagactgttaagtcgagtcgagttttcaactgacctgtgagcagaaacttcagcc 60
DB 429 GAAGAGACGTGGTAAGTGGCGGTCAGCTTTCACCTGACCTCTGGACGAGAACTTCAGCC 370
QY 61 atgaaggttaacaggaactctttctcagtcgcttggccctgttgagtcctatctggttaac 120
DB 369 ATGAAGGTAAACAGGACATCTTCTCTCAGTGGCTTGGCCCTGTTCAGTCTATCTGGTAAC 310
QY 121 actggagctgaactccctggagagagagccaaatgttaaatgaatgaatgagcc 180
DB 309 ACTGGAGCTGACTCCCTGGGAGAGAGGCGCAATGTTACAATGAACCTTAATGGATGCACC 250
QY 181 aagatatgacctgtctgtgggactgatgaaataacttatcccaatgaatgcgtgtta 240
DB 249 AAGATATAGACCTCTCTGTGGACTGATGGAATACTTATCCCAATGAATGCGTGTGA 190
QY 241 tgtttgaaggtcgaaacccagactctctcattcaaaaactctggccctgtgga 300
DB 189 TGTTTTGAAATCGGAACGCCAGACTTCTATCTCATTCATCAAAAATCTGGGCCCTTGTGA 130
QY 301 gaaccaaggtttgaaatccccatcaggctcaccgcgagcc-----tattgttga 350
DB 129 GAACCAAGGTTTGAATCCCATCAGTCCAGGCGAGGCGCTGACTGGCCTTATTGTGAA 70
QY 351 taaatgtatcgaatc 368
DB 69 TAAATGTATCTGAATATC 52

RESULT 14

US-09-760-492-57/c
: Sequence 57, Application US/09760492
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT267
: CURRENT APPLICATION NUMBER: US/09760,492
: CURRENT FILING DATE: 2001-01-16

: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 184
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 57
: LENGTH: 607
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (23)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (25)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (35)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-760-492-57

Query Match 93.7%; Score 344.8; DB 30; Length 607;

Best Local Similarity 96.8%; Pred. No. 1.5e-98;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;

QY 1 gaagagactgttaagtcgagtcgagttttcaactgacctgtgagcagaaacttcagcc 60
DB 429 GAAGAGACGTGGTAAGTGGCGGTCAGCTTTCACCTGACCTCTGGACGAGAACTTCAGCC 370
QY 61 atgaaggttaacaggaactctttctcagtcgcttggccctgttgagtcctatctggttaac 120
DB 369 ATGAAGGTAAACAGGACATCTTCTCTCAGTGGCTTGGCCCTGTTCAGTCTATCTGGTAAC 310
QY 121 actggagctgaactccctggagagagagccaaatgttaaatgaatgaatgagcc 180
DB 309 ACTGGAGCTGACTCCCTGGGAGAGAGGCGCAATGTTACAATGAACCTTAATGGATGCACC 250
QY 181 aagatatgacctgtctgtgggactgatgaaataacttatcccaatgaatgcgtgtta 240
DB 249 AAGATATAGACCTCTCTGTGGACTGATGGAATACTTATCCCAATGAATGCGTGTGA 190
QY 241 tgtttgaaggtcgaaacccagactctctcattcaaaaactctggccctgtgga 300
DB 189 TGTTTTGAAATCGGAACGCCAGACTTCTATCTCATTCATCAAAAATCTGGGCCCTTGTGA 130
QY 301 gaaccaaggtttgaaatccccatcaggctcaccgcgagcc-----tattgttga 350
DB 129 GAACCAAGGTTTGAATCCCATCAGTCCAGGCGAGGCGCTGACTGGCCTTATTGTGAA 70
QY 351 taaatgtatcgaatc 368
DB 69 TAAATGTATCTGAATATC 52

RESULT 15

US-09-647-801-11
: Sequence 11, Application US/09647801
: GENERAL INFORMATION:
: APPLICANT: Specht, Thomas
: TITLE OF INVENTION: Human Nucleic Acid Sequences which are Overexpressed
: TITLE OF INVENTION: In Normal Pancreas Tissue
: FILE REFERENCE: 51581AWOM1XX24-P
: CURRENT APPLICATION NUMBER: US/09/647,801
: CURRENT FILING DATE: 2000-10-05
: NUMBER OF SEQ ID NOS: 71
: SEQ ID NO 11
: LENGTH: 608
: TYPE: DNA
: ORGANISM: homo sapiens
US-09-647-801-11

Query Match 93.7%; Score 344.8; DB 25; Length 608;
Best Local Similarity 96.8%; Pred. No. 1.5e-98;

us-09-880-107-3847.rnps

Tue Jul 30 07:47:49 2002

	Matches	366;	Conservative	0;	Mismatches	2;	Indels	10;	Gaps	1;
QY	1	gaagagacgtggtgaagtcggtgagtttcaactgacctctggagcgagaaacttcagcc	60							
Db	204	gaagagacgtggtgaagtcggtgagtttcaactgacctctggagcgagaaacttcagcc	263							
QY	61	atgaaggtaaacagcagcatcttctctcagtcgcttggccctgttgagctctatctggtaac	120							
Db	264	atgaaggtaaacagcagcatcttctctcagtcgcttggccctgttgagctctatctggtaac	323							
QY	121	actggagctgactccctgggagagagagccaaatgttacaaatgaacttaaatgaatgcacc	180							
Db	324	actggagctgactccctgggagagagagccaaatgttacaaatgaacttaaatgaatgcacc	383							
QY	181	aagatatagaccctgtctgtgggactgatggaatacttatcccaatgaatgcgtgta	240							
Db	384	aagatatagaccctgtctgtgggactgatggaatacttatcccaatgaatgcgtgta	443							
QY	241	tgtttgaagggtcggaacgccagactctctatccctcattcaaaaatctgggcttgcctga	300							
Db	444	tgtttgaagggtcggaacgccagactctctatccctcattcaaaaatctgggcttgcctga	503							
QY	301	gaaccaaggtttgaaatcccatcaggtcacccgagagcc-----tatgtgtgaa	350							
Db	504	gaaccaaggtttgaaatcccatcaggtcacccgagagcc-----tatgtgtgaa	563							
QY	351	taaatgtatctgaatc	368							
Db	564	taaatgtatctgaatc	581							

Search completed: July 29, 2002, 19:20:58
Job time: 9799 sec

Result No.	Score	Query Match	Length	DB ID	Description
1	344.8	93.7	604	21 AAC599064	Human pancreatic c
2	344.8	93.7	608	20 AAZ41396	Human normal pancr
3	325.2	88.4	573	21 AAC04140	Human secreted pro
4	325.2	88.4	621	21 AAC04139	Human secreted pro
5	324.8	88.3	438	23 AAS91499	DNA encoding novel
6	307.8	83.6	341	9 AAN80460	Human pancreatic s
7	211.4	57.4	265	21 AAA77776	CDNA encoding huma
8	211.4	57.4	265	22 AAI28514	Colon tumour relat
9	191.8	52.1	324	23 AAS91498	DNA encoding novel

DR MPI: 2000-579444/54.
XX P-PSDB; AAB54299.
XX
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition.
XX
XX
XX Claim 1: Page 723-724; 1379pp; English.
XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
XX Sequence 604 BP: 169 A; 134 C; 143 G; 154 T; 4 other;

Query Match 93.7%; Score 344.8; DB 21; Length 604;
Best Local Similarity 96.8%; Pred. No. 8.1e-108;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;
OY 1 gaagagacgtgtaagtcggtgcagtttcaactgacctgtgacgacgaacttcagcc 60
DB 179 gaagagacgtgtaagtcggtgcagtttcaactgacctgtgacgacgaacttcagcc 238
OY 61 atgaaggttaacagcgcattcttctcagtcgcttgccctgttgagctctatctgtgaac 120
DB 239 atgaaggttaacagcgcattcttctcagtcgcttgccctgttgagctctatctgtgaac 298
OY 121 actggagctgactccctgggaagagagagccaaatgttacaataaacttaatgaatgcacc 180
DB 299 actggagctgactccctgggaagagagagccaaatgttacaataaacttaatgaatgcacc 358
OY 181 aagatatagaccctgtctgtggagctgatggaataacttaccatgaatgcgtgtta 240
DB 359 aagatatagaccctgtctgtggagctgatggaataacttaccatgaatgcgtgtta 418
OY 241 ttttttgaagtcggaacgcagacttctctcattcattcaaaaatctggcccttgcga 300
DB 419 ttttttgaatcggaaacgcagacttctctcattcattcaaaaatctggcccttgcga 478
OY 301 gaacaaagttttgaaatccccatcaggctcaccgcgagggc-----tattgtttaa 350
DB 479 gaacaaagttttgaaatccccatcaggctcaccgcgagggcctgactggccttattgttaa 538
OY 351 taatgtatctgaatc 368
DB 539 taatgtatctgaatc 556

RESULT 2
AA241396
ID AA241396 standard; cDNA; 608 BP.
XX
AC AA241396;

XX 19-JAN-2000 (first entry)
XX Human normal pancreas tissue derived cDNA 11.
XX
XX Human; pancreas; cancer; treatment; anticancer; cytostatic;
KW gene therapy; EST; expressed sequence tag; ss.
XX
XX Homo sapiens.
OS
XX DE19818598-A1.
XX
XX 21-OCT-1999.
XX
XX 19-APR-1998: 98DE-1018598.
XX
XX 19-APR-1998: 98DE-1018598.
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX WFI: 1999-592039/51.
XX
XX New nucleic acid sequences expressed in normal pancreatic tissues, and
XX derived polypeptides, for treatment of pancreatic cancer and
XX identification of therapeutic agents.
XX
XX Claim 3: Page 53; 92pp; German.
XX
XX This invention describes novel cDNA sequences (A) that are highly
XX expressed in normal human pancreatic tissue and which have anticancer
XX and cytostatic activity. (A) are used (i) for recombinant expression of
XX polypeptides (B) and (ii) to isolate complete genes. (B) are used (i) to
XX identify agents suitable for treatment of pancreatic cancer; (ii)
XX directly for treating this form of cancer (including expression from
XX gene therapy vectors) and (iii) for generation of specific antibodies.
XX (A) are identified by assembling ESTs (expressed sequence tags) from a
XX particular tissue type before comparison of expression patterns. This
XX allows a significantly longer fragment of the gene to be revealed, so
XX from different libraries may represent different parts of the same
XX unknown gene, distorting the estimated frequency of occurrence in a
XX particular tissue. AA241386-241423 represent human normal pancreatic
XX tissue derived cDNA fragments which encode the protein fragments
XX represented in AA159893-159920.
XX
XX Sequence 608 BP: 182 A; 128 C; 148 G; 150 T; 0 other;

Query Match 93.7%; Score 344.8; DB 20; Length 608;
Best Local Similarity 96.8%; Pred. No. 8.1e-108;
Matches 366; Conservative 0; Mismatches 2; Indels 10; Gaps 1;
OY 1 gaagagacgtgtaagtcggtgcagtttcaactgacctgtgacgacgaacttcagcc 60
DB 204 gaagagacgtgtaagtcggtgcagtttcaactgacctgtgacgacgaacttcagcc 263
OY 61 atgaaggttaacagcgcattcttctcagtcgcttgccctgttgagctctatctgtgaac 120
DB 264 atgaaggttaacagcgcattcttctcagtcgcttgccctgttgagctctatctgtgaac 323
OY 121 actggagctgactccctgggaagagagagccaaatgttacaataaacttaatgaatgcacc 180
DB 324 actggagctgactccctgggaagagagagccaaatgttacaataaacttaatgaatgcacc 383
OY 181 aagatatagaccctgtctgtggagctgatggaataacttaccatgaatgcgtgtta 240
DB 384 aagatatagaccctgtctgtggagctgatggaataacttaccatgaatgcgtgtta 443
OY 241 ttttttgaagtcggaacgcagacttctctcattcattcaaaaatctggcccttgcga 300
DB 444 ttttttgaatcggaaacgcagacttctctcattcattcaaaaatctggcccttgcga 503

301 gaaccaaggttttgaatcccatcaggctaccgcgagggc-----tatgtgtgaa 350
|||||
504 gaaccaaggttttgaatcccatcaggctaccgcgagggcctgactgaccttatgttga 563
|||||
351 taaatgtatctgaatc 368
|||||
564 taaatgtatctgaatc 581
|||||

RESULT 3
AAC04140
ID AAC04140 standard; cDNA: 573 BP.
XX
AC AAC04140;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 8215.
XX
KW Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 8215; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 573 BP; 145 A; 139 C; 132 G; 150 T; 7 other;

Query Match 88.4%; Score 325.2; DB 21; Length 573;
Best Local Similarity 93.2%; Pred. No. 4.1e-101;
Matches 344; Conservative 4; Mismatches 20; Indels 1; Gaps 1;

Oy 1 gaagagacgtgaatgcgagtcgagtttcaactgcacctctggacgcagaacttcagcc 60
|||||
Db 190 gaagagacgtgaatgcgagtcgagtttcaactgcacctctggacgcagaacttcagcc 249
|||||
Oy 61 atgaaggttaacaggcatctttcttcagtgcccttgccctgtgagctatcttggtaac 120
|||||

Db 250 atgaaggttaacaggcatctttcttcagtgcccttgccctgtgagctatcttggtaac 309
Oy 121 actgagctgactccctctgggaagagagccaaatgttacaatgaactaatgagtcacc 180
|||||
Db 310 actgagctgactccctctgggaagagagccaaatgttacaatgaactaatgagtcacc 369
|||||
Oy 131 aagatatagacctgtctgtgaggactatgaaatactatcccaatgaatgcgtgtta 240
|||||
Db 370 aagatatagacctgtctgtgaggactatgaaatactatcccaatgaatgcgtgtta 429
|||||
Oy 241 tgttttgaaggtcggaaacgccagactctctatccctcattcaaaaatctgggacctgtgga 300
|||||
Db 430 tgttttgaaggtcggaaacgccagactctctatccctcattcaaaaatctgggacctgtgga 489
|||||
Oy 301 -gaaccaaggttttgaatcccatcaggctaccgcgagggcctattgtgaaataatgtat 359
|||||
Db 450 ggaacnaaggttttgaatcccatcaggctaccgcgagggcctgactgcnkctwatgtwga 549
|||||
Oy 360 ctgaatctc 368
|||||
Db 550 ataaatgtc 558
|||||

RESULT 4
AAC04139
ID AAC04139 standard; cDNA: 621 BP.
XX
AC AAC04139;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 8214.
XX
KW Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 8214; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ


```

XX FH Key Location/Qualifiers
XX FT CDS 26..265
XX FT /tag= a
XX FT sig_peptide 26..94
XX FT /tag= b
XX FT mat_peptide 95..262
XX FT /tag= c
XX FT /product=PSTI
XX FT
XX FT EP267692-A.
XX PN
XX XX
XX PD 18-MAY-1988.
XX XX
XX PF 14-OCT-1987; 87EP-0309087.
XX XX
XX PR 14-MAY-1986; 86JP-0245049.
XX XX
XX PA (SHIO ) SHIONOGI SEIYAKU KK.
XX XX
XX PI Ogawa M, Matsubara K;
XX XX
XX DR WPI; 1988-134544/20.
XX DR P-PSDB; AAP80460.
XX XX
XX PT Recombinant human pancreatic secretory trypsin inhibitor
XX PT used for in vitro diagnosis and monitoring of pancreatic
XX PT diseases.
XX XX
XX PS Disclosure; ; p: English.
XX XX
XX CC Saccharomyces cerevisiae AH22 were transformed with a vector contg
XX CC this sequence. The cDNA was isolated from a pBR322 library in
XX CC E.coli K12 HB101. cDNA was prepared from mRNA obtained from human
XX CC pancreas. The library was screened firstly by a probe to codons
XX CC specifying amino acids 8-12 of mature PSTI. DNA fragments
XX CC identified in this way were then used to re-screen the library for
XX CC the full-length gene. Recombinant PSTI was obtained upon culturing
XX CC yeast transformants.
XX CC The PSTI-encoding sequence can also be chemically synthesised.
XX CC See also AAN82422.
XX XX
XX SQ Sequence 341 BP; 94 A; 76 C; 78 G; 93 T; 0 other;

Query Match 83.6%; Score 307.8; DB 9; Length 341;
Best Local Similarity 96.5%; Pred. No. 2.9e-95;
Matches 329; Conservative 0; Mismatches 2; Indels 10; Gaps 1:

Oy 36 gacctgtgagctatctgttaacactggaactgagctgactccctgggaagagagccaaatg 155
Db 1 gacctgtgagctatctgttaacactggaactgagctgactccctgggaagagagccaaatg 120
Oy 96 gacctgtgagctatctgttaacactggaactgagctgactccctgggaagagagccaaatg 155
Db 61 gacctgtgagctatctgttaacactggaactgagctgactccctgggaagagagccaaatg 120
Oy 156 ttacaatgaacttaagtgtgacccaagatatatgacctgtctgtgggaactgatggaaa 215
Db 121 ttacaatgaacttaagtgtgacccaagatatatgacctgtctgtgggaactgatggaaa 180
Oy 216 tacttatccaatgaatgcgtgtatgttttgaagtcggaacgcagacttctatct 275
Db 181 tacttatccaatgaatgcgtgtatgttttgaagtcggaacgcagacttctatct 240
Oy 276 cattcaaaaatctggcctgctgagaaaccaaagtttttgaatcccatcagggtcacccg 335
Db 241 cattcaaaaatctggcctgctgagaaaccaaagtttttgaatcccatcagggtcacccg 300
Oy 336 aggcc-----tatttgaataaatagtatctgaata 366
Db 301 aggcctgactggccttattgttgaataaataatgtatctgaata 341

```

```

RESULT 7
AAA77775
ID AAA77776 standard; cDNA: 265 BP.
XX
AC AA77776;
XX
DT 14-NOV-2000 (first entry)
XX
DE cDNA encoding human colon tumour polypeptide, SEQ ID NO:55.
XX
KW Human colon tumour polypeptide; tumour antigen; cancer; vaccine;
KW immunotherapy; diagnosis; progression; ss.
XX
OS Homo sapiens.
XX
PN WO200037643-A2.
XX
PD 29-JUN-2000.
XX
PF 23-DEC-1999; 99WO-US30909.
XX
PR 23-DEC-1998; 98US-0221298.
PR 02-JUL-1999; 99US-0347496.
PR 22-SEP-1999; 99US-0401064.
PR 19-NOV-1999; 99US-0444242.
PR 02-DEC-1999; 99US-0454150.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk J;
PI Wang T, Yugu J;
XX
DR WPI; 2000-442671/38.
XX
PT New colon tumor polypeptides used to inhibit the development of cancer,
PT especially colon cancer, and for diagnosing and monitoring the
PT progression of the cancer.
XX
PS Claim 29; Page 103; 229pp; English.
XX
CC Sequences AAA77722-A78199 represent 478 cDNAs encoding proteins or
CC portions of proteins which are associated with human colon tumours.
CC The invention also specifically discloses 8 human colon tumour proteins
CC (AAB11897-B11904). The nucleic acids, the polypeptides they encode, and
CC antigen presenting cells (APCs), preferably dendritic cells) expressing
CC such polypeptides may be used in vaccines that target tumour cells,
CC especially colon tumour cells, thereby inhibiting the development of
CC cancer. T-cells specific for the polypeptide expressed by the APC are
CC used to remove tumour cells from biological samples, especially blood or
CC fractions thereof. The sample or the isolated T-cells specific for the
CC polypeptide can then be used to inhibit cancer development. CD4+ and/or
CC CD8+ T-cells from a patient may be incubated with a polypeptide or
CC nucleic acid of the invention, or an APC expressing such a polypeptide,
CC to cause the proliferation of specific T-cells. The T-cells can be
CC cloned and then administered back to the patient to inhibit cancer
CC development. Nucleic acids encoding the polypeptides and antibodies
CC against the polypeptides may be used to determine the expression level
CC of a tumour protein of the invention, and therefore to determine whether
CC cancer cells are present. Such diagnostic methods may also be used to
CC monitor the progression of a cancer by repeating the processes at time
CC intervals, and comparing the current result to previous results. The
CC present sequence represents a cDNA encoding a human colon tumour
CC polypeptide.
XX
SQ Sequence 265 BP; 69 A; 59 C; 57 G; 75 T; 5 other;

```

```

Query Match 57.4%; Score 211.4; DB 21; Length 265;
Best Local Similarity 95.6%; Pred. No. 2.9e-62;
Matches 237; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

Oy 75 atctttcttcagtcgcttggcccttggatctatctgttaacactgagctgactcc 135

```


Db 1 atcttctctcagtcgcttgccctgttgagtcctatctggttaacactgagctgactcc 60
Qy 136 ctgggaagagagcccaaatgttacaatgaacttaattggtgacacacagatatatgacct 195
Db 61 ctgggaagagagcccaaatgttacaatgaacttaattggtgacacacagatatatgacct 120
Qy 196 gtcctggtgagactgagaaatacttaccatgaatg-cgtgttatg-ttttcaaggtc 253
Db 121 gtcctggtgagactgagaaatacttaccatgaatg-cgtgttatg-ttttcaaggtc 180
Qy 254 ggaacccagacttctatctcctcattcaaaaatctggcccttgctgagaacacagggtttt 313
Db 181 ggaacccagacttctatctcctcattcaaaaatctggcccttctgagaacacagggtttt 240
Qy 314 gaaatccc 321
Db 241 naaaatcc 248

RESULT 8
AAI28514
ID AAI28514 standard; cDNA: 265 BP.
XX AC AAI28514:
XX 12-OCT-2001 (first entry)
XX Colon tumour related cDNA sequence 11099.
XX Human; immunotherapy; diagnosis; colon cancer; colon tumour;
KW Immunogenic; gene therapy; vaccine; colonic cancer; ss.
XX Homo sapiens.
XX WO200149716-A2.
XX 12-JUL-2001.
XX 29-DEC-2000; 2000WO-US35596.
XX 30-DEC-1999; 99US-0476296.
XX 10-JAN-2000; 2000US-0480321.
XX 15-FEB-2000; 2000US-0504629.
XX 06-MAR-2000; 2000US-0519444.
XX 19-MAY-2000; 2000US-0575251.
XX 29-JUN-2000; 2000US-0609448.
XX 28-AUG-2000; 2000US-0649811.
XX (CORI-) CORIXA CORP.
XX Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stoik JA;
PI King GE, Wang T, Jiang Y;
XX WPI; 2001-441847/47.
XX Colon tumor associated proteins and nucleic acids useful for the
PT prevention, diagnosis and treatment of colonic cancer -
XX Claim 25; Page 145; 472pp; English.

XX The present invention describes colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be
CC used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the
CC activity of TCAPs by expressing inactive proteins or to supplement the
CC patients own production of them. Additionally, (II) may be used to
CC produce the TCAP proteins, by inserting the nucleic acids into a host
CC cell culturing the cell to express the protein. (II) and its

CC complementary sequences may also be used as DNA probes in diagnostic
CC polymerase chain reaction (PCR) and hybridisation assays to detect and
CC quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. (I) may
CC also be used as antigens in the production of antibodies against TCAPs
CC and in assays to identify modulators of TCAP expression and activity.
CC Anti-(I) antibodies and antagonists may also be used to down regulate
CC TCAP expression and activity. The anti-(I) antibodies may also be used
CC as diagnostic agents for detecting the presence of TCAPs in samples
CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AAI28460 to AAI29512
CC and AAI24494 to AAI24523 represent nucleotide and amino acid sequences
CC given in the exemplification of the present invention.
XX Sequence 265 BP: 69 A: 59 C: 57 G: 75 T: 5 other:
SQ

Query Match 57.4%; Score 211.4; DB 22; Length 265;
Best Local Similarity 95.8%; Pred. No. 2.9e-62;
Matches 237; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
Qy 76 atcttctctcagtcgcttgccctgttgagtcctatctggttaacactgagctgactcc 135
Db 1 atcttctctcagtcgcttgccctgttgagtcctatctggttaacactgagctgactcc 60
Qy 136 ctgggaagagagcccaaatgttacaatgaacttaattggtgacacacagatatatgacct 195
Db 61 ctgggaagagagcccaaatgttacaatgaacttaattggtgacacacagatatatgacct 120
Qy 136 gtcctggtgagactgagaaatacttaccatgaatg-cgtgttatg-ttttcaaggtc 253
Db 121 gtcctggtgagactgagaaatacttaccatgaatg-cgtgttatg-ttttcaaggtc 180
Qy 254 ggaacccagacttctatctcctcattcaaaaatctggcccttgctgagaacacagggtttt 313
Db 181 ggaacccagacttctatctcctcattcaaaaatctggcccttctgagaacacagggtttt 240
Qy 314 gaaatccc 321
Db 241 naaaatcc 248

RESULT 9
AAS9149E
ID AAS91498 standard; cDNA: 324 BP.
XX AC AAS91498;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #27302.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 25-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX DrTanJC RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-FSD3; ABG27311.
XX New isolated polynucleotide and encoded polypeptides, useful in
PI

PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 1; SEQ ID No 27302; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 324 BP; 76 A; 68 C; 87 G; 93 T; 0 other;

Query Match 52.1%; Score 191.8; DB 23; Length 324;
 Best Local Similarity 99.0%; Pred. No. 1.6e-55;
 Matches 193; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 61 atgaaggttaacagcctcttcttcagtcgttgccctgttgagtcctatcgtgtaac 120
 DB 1 atgaaggttaacagcctcttcttcagtcgttgccctgttgagtcctatcgtgtaac 60
 QY 121 actggagctgactccctcctggagagagagccaaatgttacaatgaacttaagtgaacc 180
 DB 61 actggagctgactccctcctggagagagagccaaatgttacaatgaacttaagtgaacc 120
 QY 181 aagatatagacctgtctgtggagctgatgaaatcactatcccaatgaatgcgtgta 240
 DB 121 aagatatagacctgtctgtggagctgatgaaatcactatcccaatgaatgcgtgta 180
 QY 241 tgtttgaaagtcgg 255
 DB 181 tgtttgaaagtcgg 195

RESULT 10
 AAT84523
 ID AAT84523 standard; DNA; 238 BP.
 AC AAT84523;

XX 02-DEC-1997 (first entry)

XX hpSTI.M2 gene for baculoviral expression.

XX Subtilisin inhibitor; protease inhibitor; PSTI; hpSTI.M2; human;
 KW pancreatic secretory trypsin inhibitor; enzyme engineering;
 KW protein engineering; baculovirus; detergent; ss.

XX Chimeric Homo sapiens;
 OS Chimeric synthetic.

XX Key Location/Qualifiers
 FT CDS 1..228
 FT /*Cag- a

PN WO9715670-A1.
 XX 01-MAY-1997.
 XX 25-OCT-1996; 96WO-US17153.
 XX 25-OCT-1995; 95US-0548186.
 XX (ARRI-) ARRIS PHARM CORP.
 PA Estell DA, Fisher J, Hartman C, McGrath M, Schmidt B;
 PI WPI; 1997-259023/23.
 DR P-PSDB; AAW26361.
 XX Modified subtilisin inhibitor - having altered reactive site
 PT resulting in decreased immunogenicity, used in detergent
 PT composition(s)
 PS Example 1; Fig 8; 69pp; English.
 XX This DNA sequence encodes a polypeptide (AAW26361) comprising human
 CC pancreatic secretory trypsin inhibitor (PSTI), modified to include a
 CC C-terminal M2 tag. The sequence is suitable for baculovirus
 CC expression. Novel modified subtilisin inhibitors are claimed in
 CC in which a human standard mechanism inhibitor, such as PSTI, has at
 CC least one amino acid (aa) of the reactive site substituted with
 CC different aa so that its dissociation constant for inhibition (Ki),
 CC for at least one subtilisin, is reduced by at least a factor of
 CC 100. Also claimed are: a nucleic acid (I) encoding the subtilisin
 CC inhibitor; expression vectors containing (I); and host cells
 CC containing (I). The subtilisin inhibitors (see AAW26362-64) combine
 CC the low allergenicity of human standard mechanism inhibitors, which
 CC are not specific for subtilisin, and the high affinity of non-human
 CC subtilisin inhibitors. They form a complex with subtilisin, so as
 CC to avoid problems of allergenicity in detergent formulations, and
 CC can also be used in affinity purification and (diagnostic)
 CC quantification of subtilisin.
 XX Sequence 238 BP; 68 A; 58 C; 61 G; 51 T; 0 other;

Query Match 47.2%; Score 173.8; DB 18; Length 238;
 Best Local Similarity 96.2%; Pred. No. 2.1e-49;
 Matches 178; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 113 ctggttaacactggagctgactccctcctggagagagagccaaatgttacaatgaacttaag 172
 DB 11 ccggcaacaactggagctgactccctcctggagagagagccaaatgttacaatgaacttaag 70
 QY 173 gatgcaccaagatatatgacctgtctgtggagctgatgaaatcactatcccaatgaat 232
 DB 71 gatgcaccaagatatatgacctgtctgtggagctgatgaaatcactatcccaatgaat 130
 QY 233 gcgtgttatgttttgaaggctcggaacgccagactctatcctcattcaaaaatcgtggcc 292
 DB 131 gcgtgttatgttttgaaggctcggaacgccagactctatcctcattcaaaaatcgtggcc 190
 QY 293 ctgac 297
 DB 131 ctgac 195

RESULT 11
 AAN81535
 ID AAN81535 standard; DNA; 179 BP.

XX AAN81535;

XX 07-DEC-1999 (first entry)

XX Human pancreatic secretory trypsin inhibitor synthetic gene.

XX Yoshiida N, Kikuchi N, Teraoka H, Shin M;
XX WPI: 1990-024464/04
XX P-PSDB: AAR03727.
XX Human PSTII modified at arginine 42 and/or 44
XX by substitution with glutamine and/or serine to allow
XX trypsin-inhibitory activity.
XX Disclosure: fig 1: 25pp; English.
XX This modified human pancreatic secretory trypsin inhibitor
XX differs from natural human PSTI in that bases 124-26 that
XX usually code for Arg are replaced with Gln. The modified
XX PSTI allows a sustained trypsin inhibition effect useful
XX in the clinical treatment of acute pancreatitis. See also
XX AAQ01494, AAQ01493, AAR03727, AAR03728 and AAR03729.
XX Sequence 171 BP; 53 A; 35 C; 37 G; 46 T; 0 other;
SQ

Query Match 45.2%; Score 166.2; DB 11; Length 171;
Best Local Similarity 98.2%; Pred. No. 7.2e-47;
Matches 168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 130 gactccctgggaagagagcccaaatgttacaatgaacttaattggaatggaacccaagatat 189
DB 1 gactccctgggaagagagcccaaatgttacaatgaacttaattggaatggaacccaagatat 60

QY 190 gaccctgtgtgtgggactgatggaataacttaccatgaatgcgtgtgtgtttgaa 249
DB 61 gaccctgtgtgtgggactgatggaataacttaccatgaatgcgtgtgtttgaa 120

QY 250 ggtcgaaacgcagactctctatccctcattcaaaaatctgggacctgtgta 300
DB 121 aatcagaacgcagactctctatccctcattcaaaaatctgggacctgtgta 171

RESULT 14
AAAT7777/c
ID AAA77777 standard; cDNA: 420 BP.
XX
XX AAA77777;
XX
XX 14-NOV-2000 (first entry)
XX cDNA encoding human colon tumour polypeptide, SEQ ID NO:56.
XX Human colon tumour polypeptide; tumour antigen; cancer; vaccine;
XX immunotherapy; diagnosis; progression; ss.
XX Homo sapiens.
XX OS
XX WO200037643-A2.
XX
XX 29-JUN-2000.
XX
XX 23-DEC-1999; 99WO-US30909.
XX
XX 23-DEC-1998; 98US-0221298.
XX 02-JUL-1999; 99US-0347496.
XX 22-SEP-1999; 99US-0401064.
XX 19-NOV-1999; 99US-0444242.
XX 02-DEC-1999; 99US-0454150.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stoik J;
XX Wang T, Yuqiu J;
XX WPI: 2000-442671/38.
XX

PT New colon-tumor polypeptides used to inhibit the development of cancer,
PT especially colon cancer, and for diagnosing and monitoring the
PT progression of the cancer -
XX Claim 29: Page 103; 229pp; English.
XX Sequences AAA77722-A78199 represent 478 cDNAs encoding proteins or
XX portions of proteins which are associated with human colon tumours.
XX The invention also specifically discloses 8 human colon tumour proteins
XX (AAB1897-B11904). The nucleic acids, the polypeptides they encode, and
XX antigen presenting cells (APCs), preferably dendritic cells) expressing
XX such polypeptides may be used in vaccines that target tumour cells,
XX especially colon tumour cells, thereby inhibiting the development of
XX cancer. T-cells specific for the polypeptide expressed by the APC are
XX used to remove tumour cells from biological samples, especially blood or
XX fractions thereof. The sample or the isolated T-cells specific for the
XX polypeptide can then be used to inhibit cancer development. CD4+ and/or
XX CD8+ T-cells from a patient may be incubated with a polypeptide or
XX nucleic acid of the invention, or an APC expressing such a polypeptide,
XX to cause the proliferation of specific T-cells. The T-cells can be
XX cloned and then administered back to the patient to inhibit cancer
XX development. Nucleic acids encoding the polypeptides and antibodies
XX against the polypeptides may be used to determine the expression level
XX of a tumour protein of the invention, and therefore to determine whether
XX cancer cells are present. Such diagnostic methods may also be used to
XX monitor the progression of a cancer by repeating the processes at time
XX intervals, and comparing the current result to previous results. The
XX present sequence represents a cDNA encoding a human colon tumour
XX polypeptide.
XX Sequence 420 BP; 103 A; 107 C; 88 G; 111 T; 11 other;
SQ

Query Match 44.1%; Score 162.2; DB 21; Length 420;
Best Local Similarity 96.2%; Pred. No. 2.6e-45;
Matches 175; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 160 aatgaacttaattggaatggaacccaagatatgacctgtgtgtggactgatggaataac 218
DB 230 AANGAACTTAATGGATGACCCCAAGATATATGACCCNGTCTGTGGGACTCATGGAATAC 141

QY 219 ttatcccaatgaatgcgtgtgttatgtttgaaggtcggaacccagacctctatctcat 278
DB 140 TTATCCCAATCAATGCGTGTATGTTTGAATAATCGGAACGCCAGACTTCTATCTCAT 81

QY 279 tcaaaaatctgggacctgtgtgtagaaccaaggttttgaatcccatcaggtcaccgcgag 338
DB 20 TCAAAAATCTGGGCTTGTCTGTGAGAACCAAGGTTTGAATCCCATCAGGTACCGCGAGG 21

QY 339 AC 340
DB 20 AC 19

RESULT 15
AA128515/c
ID AA128515 standard; cDNA: 420 BP.
XX
XX AA128515;
XX
XX 12-OCT-2001 (first entry)
XX Colon tumour related cDNA sequence 11186.
XX Human immunotherapy; diagnosis; colon cancer; colon tumour;
XX immunogenic; gene therapy; vaccine; colonic cancer; ss.
XX Homo sapiens.
XX OS
XX WO200149716-A2.
XX
XX 12-JUL-2001.
XX

Job time: 5158 sec

PF XX 29-DEC-2000: 2000WO-US35596.
PR 30-DEC-1999: 99US-0476296.
PR 10-JAN-2000: 2000US-0480321.
PR 15-FEB-2000: 2000US-0504629.
PR 06-MAR-2000: 2000US-0519444.
PR 19-MAY-2000: 2000US-0575251.
PR 29-JUN-2000: 2000US-0609448.
PR 28-AUG-2000: 2000US-0649811.
XX PA (CORI-) CORIXA CORP.
XX
XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;
XX
XX WPI: 2001-441847/47.
XX
XX Colon tumor associated proteins and nucleic acids useful for the
PT prevention, diagnosis and treatment of colonic cancer -
XX
XX Claim 25; Page 145; 472pp; English.
XX
XX The present invention describes colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be
CC used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the
CC activity of TCAPs by expressing inactive proteins or to supplement the
CC patients own production of them. Additionally, (II) may be used to
CC produce the TCAP proteins, by inserting the nucleic acids into a host
CC cell culturing the cell to express the protein. (II) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC polymerase chain reaction (PCR) and hybridisation assays to detect and
CC quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. (I) may
CC also be used as antigens in the production of antibodies against TCAPs
CC and in assays to identify modulators of TCAP expression and activity.
CC Anti-(I) antibodies and antagonists may also be used to down regulate
CC TCAP expression and activity. The anti-(I) antibodies may also be used
CC as diagnostic agents for detecting the presence of TCAPs in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)). AA128460 to AA129512
CC and AA24494 to AA24523 represent nucleotide and amino acid sequences
CC given in the exemplification of the present invention.
XX
XX Sequence 420 BP; 103 A; 107 C; 88 G; 111 T; 11 Other;
SQ

Query Match 44.1%; Score 162.2; DB 22; Length 420;
Best Local Similarity 96.2%; Pred. NO. 2.6e-45;
Matches 175; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Qy 160 aatgaactaatggtgca-ccaatatatgacccctgtctgtggagctgatgaaatac 218
Db 200 AANGAACTTAATGATGACACCAAGATATATGACCCNGTCTGTGGACATGAGGAATAC 141
Qy 219 ttatcccaatgaatcggtgttatgtttgaaagtcggaacgccagacttctatcctcat 278
Db 140 TTATCCCAATGAATGCGTGTATGTTTGAATAATCGGAACGCCAGACTTCTATCNTCAT 81
Qy 279 tcaaaaatctggccttctgtagaaccaggtttgaaatcccatcaggtcacccgaggg 338
Db 80 TCAAAATCTGGGCGCTTGTCTGAGACCAAGGTTTGAATCCATCCATCAGGTCACCGGAGG 21
Qy 339 cc 340
Db 20 AC 19

Search completed: July 29, 2002, 18:01:12

GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 17:20:29 ; Search time 394.41 Seconds
(without alignments)
2492.057 Million cell updates/sec

Title: US-09-880-107-2492

Perfect score: 466

Sequence: 1 atcaagaaacatagatttc.....gggtaagggtggatgtctt 466

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1333638 seqs, 1054605264 residues

Total number of hits satisfying chosen parameters: 2667276

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA, New:*

1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*

2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*

3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*

4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*

5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*

6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*

7: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*

8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
c 1	397	85.2	777	1	PCT-US02-18947-1279
c 2	397	85.2	777	7	US-10-171-311-229
c 3	397	85.2	777	7	US-10-172-118-1279
c 4	343.8	73.8	477	7	US-10-102-524-299
c 5	334.4	71.8	474	7	US-10-040-862-205
c 6	334.4	71.8	474	7	US-10-040-862-646
c 7	334.4	71.8	474	7	US-10-040-862-4647
c 8	323.4	69.4	453	7	US-10-040-862-248
c 9	323.4	69.4	453	7	US-10-040-862-5273
c 10	45.8	9.8	1274	7	US-10-137-036-34
c 11	44.2	9.5	442	5	US-09-620-393B-2769
c 12	44.2	9.5	501	5	US-09-620-393B-4443
c 13	43.4	9.3	243	5	US-09-975-254-22419
c 14	42.6	9.1	552	5	US-09-620-393B-1362
c 15	42.2	9.1	469	5	US-09-620-393B-868
c 16	42.2	9.1	469	5	US-09-935-625-18893
c 17	41	8.8	455	5	US-09-935-625-18645
c 18	41	8.8	595	5	US-09-620-393B-1166
c 19	41	8.8	596	5	US-09-620-393B-6996
c 20	39.8	8.5	554	5	US-09-886-492-17764
c 21	39.4	8.5	452	5	US-09-935-625-21348
c 22	39.4	8.5	461	5	US-09-620-393B-6747
c 23	39	8.4	461	5	US-09-620-393B-5212
c 24	39	8.4	461	5	US-09-935-625-2477
c 25	39	8.4	461	5	US-09-935-625-6719

c 26	39	8.4	461	5	US-09-935-625-11030	Sequence 11030, A
c 27	39	8.4	461	5	US-09-935-625-15719	Sequence 15719, A
c 28	39	8.4	461	5	US-09-935-625-24218	Sequence 24218, A
c 29	39	8.4	461	5	US-09-935-625-27443	Sequence 27443, A
c 30	39	8.4	492	5	US-09-620-393B-4294	Sequence 4294, Ap
c 31	38.8	8.3	465	5	US-09-620-393B-9008	Sequence 9008, Ap
c 32	38.8	8.3	1752	5	US-09-886-432-16050	Sequence 16050, A
c 33	38.2	8.2	251	5	US-09-975-254-890	Sequence 890, App
c 34	37.8	8.1	479	5	US-09-620-393B-8483	Sequence 8483, Ap
c 35	37.8	8.1	485	5	US-09-620-393B-7814	Sequence 7814, Ap
c 36	37.4	8.0	256	5	US-09-975-254-17947	Sequence 17947, A
c 37	37.2	8.0	221	5	US-09-919-002-652	Sequence 652, App
c 38	37	7.9	3083	7	US-10-137-036-1	Sequence 1, Appli
c 39	37	7.9	3084	7	US-10-137-036-143	Sequence 143, App
c 40	36.8	7.9	399	5	US-09-918-995-7727	Sequence 7727, Ap
c 41	36.6	7.9	239	5	US-09-975-254-1973	Sequence 1973, Ap
c 42	36.6	7.9	265	5	US-09-975-254-13012	Sequence 13012, A
c 43	35.8	7.7	234	5	US-09-975-254-19108	Sequence 19108, A
c 44	35.8	7.7	261	5	US-09-975-254-18104	Sequence 18104, A
c 45	35.6	7.6	552	1	PCT-US02-19457-108	Sequence 108, App

ALIGNMENTS

RESULT 1

PCT-US02-18947-1279/C

: Sequence 1279, Application PC/TUS0218947

: GENERAL INFORMATION:

: APPLICANT: Rosetta Inpharmatics

: TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

: FILE REFERENCE: 9301-175-228

: CURRENT APPLICATION NUMBER: PCT/US02/18947

: CURRENT FILING DATE: 2002-06-14

: PRIOR APPLICATION NUMBER: 60/380,770

: PRIOR FILING DATE: 2002-05-14

: NUMBER OF SEQ ID NOS: 2699

: SEQ ID NO 1279

: LENGTH: 777

: TYPE: DNA

: ORGANISM: Homo sapiens

: PUBLICATION INFORMATION:

: DATABASE ACCESSION NUMBER: NM_006398

: DATABASE ENTRY DATE: 2001-06-18

PCT-US02-18947-1279

Query Match	85.2%;	Score 397;	DB 1;	Length 777;
Best Local Similarity	98.0%;	Pred. No. 4.3e-111;		
Matches 444;	Conservative	0;	Mismatches 5;	Indels 4;
				Gaps 4;
Qy	1	atcaagaaacatagatttcgggcaataatacttcatctaccatcccccacaaatctta	60	
Db	631	NYCAAGAAACATAGAGTT-GGGCAATATACTTCTACCCATCCCAAAATCTTA	633	
Qy	61	ctctactcatctcatctcatttaattttgggaatcatcagaagatgttctcgttgagta	120	
Db	632	CTCTACTCATCTCATTTCTAATAATTTTGGAAATCATCAGAAGATGTTCTCGTGAGTA	573	
Qy	121	aaagattaaaagaaataagctttttgacccctgccacacccccatgccagggtgttcac	180	
Db	572	AGAGATTAAAGAAATAGCTTTTIGACCCCTGCCAACACCCCATCCCGGGTGTTCAC	513	
Qy	161	cttccaatacaatdacaatgcaggagagtaagtgtccctttctgaatccgtaatctgcc	240	
Db	512	CTCCAATACAATAAGATGCCAGGAGAGTAAGTTGCCCTTCTGTATGCCGTAAATCTGCC	453	
Qy	241	atcatcttcccattctccagttctcttccattgcaatgcacaaatctgggtctcagggat	300	
Db	452	ATCATCTTCCCATCTTCCAGTCT-CTTCCATTGCAAGTCACAAATCTGGGTCTCAGGGAT	394	
Qy	301	tatacccgctttagttctgatgattgttttcaactgtgccactgagctgagaccttcgac	360	

Db	393	TATACCGCTT	AGATCATTTG	TGCTTTCAC	TGTGCCACT	GAGCTGGAC	CTTCGGAC	334
Qy	361	ctggagagg	tgctctttt	tgctcat	ccactgac	tccacaag	aaacagg	420
Db	333	CT-GGAGGAG	TGCTCTTT	TGCTCAT	CACCTTGAC	TCCACAAG	AAACAGG	275
Qy	421	catcact	ggggcttc	accacttc	agggtaag			453
Db	274	CATCACT	-GGGCTT	CACCACTT	TCAGGTTAAG			243

```

RESULT      2
US-10-171-311-229/c
: Sequence 229, Application US/10171311
: GENERAL INFORMATION:
: APPLICANT: Schlegel, Robert
: APPLICANT: Chen, Yan
: APPLICANT: Zhao, Xumei
: APPLICANT: Monahan, John
: APPLICANT: Kamatkar, Shubhangi
: APPLICANT: Glatt, Karen
: APPLICANT: Gannavarapu, Manjula
: APPLICANT: Hoersh, Sebastian
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
: TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
: TITLE OF INVENTION: OF CERVICAL CANCER
: FILE REFERENCE: MRI-035
: CURRENT APPLICATION NUMBER: US/10/171.311
: CURRENT FILING DATE: 2002-06-12
: PRIOR APPLICATION NUMBER: US 60/298,159
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,155
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/335,936
: PRIOR FILING DATE: 2001-11-14
: NUMBER OF SEQ ID NOS: 238
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 229
: LENGTH: 777
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-171-311-229

```

Query Match	85.2%	Score	397;	DB	7:	Length	777;
Best Local Similarity	98.0%;	Pred.	No. 4.3e-111;				
Matches	444;	Conservative	0;	Mismatches	4;	Gaps	4;
Oy	1	atcaagaacaatagatttcgagggaataatactcattcatcccatccccccaacattta	60				
Db	691	ATCAAGAACAATAGATT-GGGCANATACTTACTCTACCCATGCCACCCANAAICTTA	633				
Oy	61	ctctactcatctcattctcataaattttggaaatcatcgaaagtgttgttgtagta	120				
Db	632	CTCTACTCATCTCATTCTCATTAATTTTGGCAATCATCAGAAGATGTGTCGTGAGTA	573				
Oy	121	agagattaaaaaataaagctttttgacctctgcacaaccccattgcccaagggtgatcac	180				
Db	572	AGAGATTAAAAGAAATAAGCTTTTTGACCCCTGCACAACCCCCATGCCCAAGGTGTGTAC	513				
Oy	181	cctcctaatacaataacatgccagggaagtagtctgcctttctgatgcqgtaactctgcc	240				
Db	512	CTTCCAATACAAATAAGATGCCAGGAAGTAGTAAGTGTGCCCTTCTGTATGCCGTAATCTGCC	453				
Oy	241	atcatotttcccatctttccagtcctcttccatttccaagtccaacatctgggtctcagggat	300				
Db	452	ATCATCTTCCCATCTTCCAGTCT-CTTTCCCATTGCAAAGTCAACAICTGGGTCTCAGGGAT	394				
Oy	301	tatacccgcttaagtcctgaatcattgctttaccitgtgcacactgaactgaactctccac	360				
Db	393	TATACC CGCTT TAGTCTCGATCATTTGCTTTAC TTGTGTG CAC T GAG CTG GACCTT GC AC	334				
Oy	361	ctggagagaggtgcctctttgtcctcatcaccttgactccacagaagaacaaggcagctct	420				

```

Db      333  CT-GGAGGAGGTGCCTCTTTTGCCTCATCACTCACTCCACAGAAACAGGGCAGCTCCT 275
|| |||||
Qy      421  catcactggggcttcaccacttcctcaggggtaag 453
||| |||||
Db      274  CATCACT-GGGCTTCACCACCTTTTCAGGTAAG 243
||| |||||

RESULT 3
US-10-172-118-1279/c
: Sequence 1279, Application US/10172118
: GENERAL INFORMATION:
: APPLICANT: Dai, Hongyue
: APPLICANT: He, Yudong
: APPLICANT: Linsley, Peter
: APPLICANT: Mao, Mao
: APPLICANT: Roberts, Chris
: APPLICANT: Van 't veer, Laura
: APPLICANT: Van de Vijver, Marc
: APPLICANT: Bernards, Rene
: TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
: FILE REFERENCE: 9301-175-999
: CURRENT APPLICATION NUMBER: US/10/172,118
: CURRENT FILING DATE: 2002-06-14
: PRIOR APPLICATION NUMBER: 60/380,770
: PRIOR FILING DATE: 2002-05-14
: NUMBER OF SEQ ID NOS: 2699
: SEQ ID: NO 1279
: LENGTH: 777
: TYPE: DNA
: ORGANISM: Homo sapiens
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: NM_006398
: DATABASE ENTRY DATE: 2001-06-18
: US-10-172-118-1279

```

Query Match	85.2%	Score 397	DB 7	Length 777
Best Local Similarity	98.0%	Pred. No. 4.3e-111		
Matches 444	Conservative	0	Mismatches 5	Indels 4
				Gaps
Oy	1	atcaagaacatagagttcgggcaataatactctacatctaccatccaccacaaatctta	60	
Db	691	ATCAAGAAACATAGAGTT -GGGCAATATATCTTCATCTACCATCCACCCAAATCTTA	633	
Oy	61	ctctactcatctcattctcatttaattttggggaatcatcagaagatgtgttcgttgagta	120	
Db	632	CTCTACTCATCTCAATCTCATTAATTTTGGGAAATCATCAGAAGATGTGTTCGTTGAGTA	573	
Oy	121	agagattaaaagaataaagctltttgacccctcgccaaccccccatgccagaggtgggtcac	180	
Db	572	AGAGATTAAAGAANAATAGCTTTTGTACCCCTGCCAACCCCATCCCCAGGGTGGTCAC	513	
Oy	181	ctccaatacaataacatgccaggaagagtaagtgtgccctttctgtaigcgttaactgcc	240	
Db	512	CCTCAATACATAAGATGTGCCAGGAAGAGTAAGTTGGCTTTCTGTAGTCGTAATCTGCC	453	
Oy	241	atcattctcccatcttccagctctcttccatgcaagtacaaatctgggtctcaggagat	300	
Db	452	ATCATCTTCCCATCTTCCAGTCT -CTTTCATTTGCAAGTCACAAATCTGGTCTCAGGGAT	394	
Oy	301	tatacccgctctagctctgatcatgtctttcaacttggccactgagctggagaccttcgac	360	
Db	393	TATACCCGCTCTTAGTCTCGATCATGTCTTTCACPTTGTGCCACTGAGCTGGACCTTCGCCAC	334	
Oy	361	ctggagagagtgctctttgctctcatcactgactccacaagaacaaagggcagctcct	420	
Db	333	CT -GGAGGAGGTGGCTCTTTGCCCTATCACCCTGACTCTCACAGAAACAGGGCAGCTCCT	275	
Oy	421	catcactggggcttcaaccactttcagggggtaag	453	
Db	274	CATCACT -GGGCTTCACCACCTTTCAGGGTAAGG	243	

RESULT 4
US-10-102-524-299/c
: Sequence 299, Application US/10102524
: GENERAL INFORMATION:
: APPLICANT: Algate, Paul A.
: APPLICANT: Mannion, Jane
: APPLICANT: Gaiger, Alexander
: APPLICANT: Mannion, Jane
: APPLICANT: Gordon, Brian
: APPLICANT: Harlocker, Susan L.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
: TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
: FILE REFERENCE: 210121.572
: CURRENT APPLICATION NUMBER: US/10/102,524
: CURRENT FILING DATE: 2002-03-19
: NUMBER OF SEQ ID NOS: 1863
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 299
: LENGTH: 477
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-102-524-299

Query Match 73.8%; Score 343.8; DB 7; Length 477;
Best Local Similarity 98.7%; Pred. No. 6.6e-95;
Matches 378; Conservative 0; Mismatches 2; Indels 3; Gaps 3:

QY 1 atcaagaacacatagattcggccaatactctaccatccccccaaatctta 60
DB 380 ATCAAGAACAACATAGATT -GGCAATATACTTCTACCTACCATCCACCAAAATCTTA 322
QY 61 ctctactcatctctctcatttaatttgggaaatcatcagaagatgttctgtagta 120
DB 321 CTCTACTCATCTCATTTCTCATTAATTTGGGAAATCATCAGAAGATGTGTCTGTGAGTA 262
QY 121 agagattaaagaataagcttttgcacctcgccacaccccatgccagggtgtcac 180
DB 261 ACAGATTAAAGAAATAAGCTTTTGACCCCTGCCAACACCCCATCCACGGGTGGTCAC 202
QY 181 ctctcaatacataacatgccaggaagaagtaagtgccttctgatgcgtaactgcc 240
DB 201 CTTCAATACATAAGATGCCAGGAAGATGAAGTTGCCCTTTCTGATGCCGTAACTGCC 142
QY 241 atcatcttcccatcttccagctctccttctccattgcaagtcacaatctgggtctcaggat 300
DB 141 ATCATCTTCCCATCTTCCAGTCT -CTTTCATTCGAAGTCACAACTCTGGGTCTCAGGGAT 83
QY 301 tataccgtcttagctcgatcattgtctcacttggccactgagctgagaccttcgcac 360
DB 82 TATACCCGCTTAGTCTCGATCATGCTTTCACITGTGCCACTGAGCTGGACCTTCGCAC 23
QY 361 ctgggaggaggtgcctctttgcc 383
DB 22 CT-GGAGGAGGTGCCTCTTTGCC 1

RESULT 5
US-10-040-862-205/c
: Sequence 205, Application US/10040862
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Algate, Paul A.
: APPLICANT: Mannion, Jane
: APPLICANT: Retter, Marc
: APPLICANT: Corixa Corporation
: TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
: TITLE OF INVENTION: Hematological Malignancies
: FILE REFERENCE: 014058-013520US
: CURRENT APPLICATION NUMBER: US/10/040,862
: CURRENT FILING DATE: 2001-11-06
: PRIOR APPLICATION NUMBER: US 60/186,126
: PRIOR FILING DATE: 2000-03-01


```

: APPLICANT: Retter, Marc
: APPLICANT: Corixa Corporation
: TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
: FILE REFERENCE: 014058-013520US
: CURRENT APPLICATION NUMBER: US/10/040,862
: CURRENT FILING DATE: 2001-11-06
: PRIOR APPLICATION NUMBER: US 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: US 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: US 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: US 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: US 60/200,779
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: US 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: US 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: US 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: US 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: US 09/796,692
: PRIOR FILING DATE: 2001-03-01
: NUMBER OF SEQ ID NOS: 10467
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 646
: LENGTH: 474
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(474)
: OTHER INFORMATION: n = A,T,C or G
US-10-040-862-646

Query Match 71.8%; Score 334.4; DB 7; Length 474;
Best Local Similarity 99.2%; Pred. No. 4.9e-92;
Matches 357; Conservative 0; Mismatches 1; Indels 2; Gaps 2:

Qy 1 atcaagaacatagagtcgggcaataactctcattctcaccatcccccacaaatctta 60
Db 358 ATCAAGAACAATAGAGTT-GGGCAATATCTTCACTTACCCTCCACCCCAAACTTAA 300
Qy 61 ctctactctcattctcatttaatttgggaatcatcagaagatggttcgtgagta 120
Db 299 CTCCTACTCACTCTCACTTCACTTAAATTTGGGAAATCATCAGAAGATGTTGTCGTGAGTA 240
Qy 121 agagataaagaataaagctttttgacccctgccaacacccatgccagggtggtcac 180
Db 239 AGAGATTAAGAATAAAGCTTTTGTATCCCTGCCAACACCCCATGCCAGGGGTGTCAC 180
Qy 181 cctccatacaataacatccagggaagataagttgcccctttctgatcccgtaactcgtcc 240
Db 179 CCTCCAATACAATAACATGCCAGGAAGATGAGTTGCCCTTTCTGATGCCGTAATCTGCC 120
Qy 241 atcatctccatctccagttctccttccattgcaagtcacaactcgtggtctcaggat 300
Db 119 ATCATCTCCCATCTTCCAGTCT-CTTTCATTCGAATGCAAACTCGGCTCAGGAT 61
Qy 301 tataccgctttagtctcgatcatctgttcactgtgcccactgagctgagcttcgcac 360
Db 60 TATACCGTCTTAGTCTCGATCATTTGCTTCACTTGTGCCACTGAGCTGGACCTTCGCAC 1

```

```

RESULT 7
US-10-040-862-4647/c
: Sequence 4647, Application US/10040862
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Algate, Paul A.
: APPLICANT: Mannion, Jane
: APPLICANT: Retter, Marc
: APPLICANT: Corixa Corporation
: TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
: FILE REFERENCE: 014058-013520US
: CURRENT APPLICATION NUMBER: US/10/040,862
: CURRENT FILING DATE: 2001-11-06
: PRIOR APPLICATION NUMBER: US 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: US 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: US 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: US 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: US 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: US 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: US 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: US 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: US 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: US 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: US 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: US 60/223,378
: PRIOR FILING DATE: 2000-08-07
: PRIOR APPLICATION NUMBER: US 09/796,692
: PRIOR FILING DATE: 2001-03-01
: NUMBER OF SEQ ID NOS: 10467
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4647
: LENGTH: 474
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (446)
: OTHER INFORMATION: n=A,T,C or G
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (463)
: OTHER INFORMATION: n=A,T,C or G
US-10-040-862-4647

Query Match 71.8%; Score 334.4; DB 7; Length 474;
Best Local Similarity 99.2%; Pred. No. 4.9e-92;
Matches 357; Conservative 0; Mismatches 1; Indels 2; Gaps 2:

Qy 1 atcaagaacatagagtcgggcaataactctcattctcaccatcccccacaaatctta 60
Db 358 ATCAAGAACAATAGAGTT-GGGCAATATCTTCACTTACCCTCCACCCCAAACTTAA 300
Qy 61 ctctactctcattctcatttaatttgggaatcatcagaagatggttcgtgagta 120
Db 299 CTCCTACTCACTCTCACTTCACTTAAATTTGGGAAATCATCAGAAGATGTTGTCGTGAGTA 240
Qy 121 agagataaagaataaagctttttgacccctgccaacacccatgccagggtggtcac 180

```

Db 239 AGAGATTAAAGAAATAAGCTTTTGTATCCCTGCGCAACACCCCATGCCAGGGTGGTCAC 180
Qy 181 cctccatacaataacatccaggagagtaagtgtgcccctttctgtatgccgttaattctgcc 240
Db 179 CTCTCAATACATAACATCCAGGAAGAGTAAGTTGCCCTTTCTGATGCCGTAATCTGCC 120
Qy 241 atcatcttccatcttccagctctcttccatgcacaaactcgggtctcaaggat 300
Db 119 ATATCTCTCCCATCTTCCAGTCT-CTTCCATGCGAAGTCACAATCTGGGTCTCAGGAT 61
Qy 301 tataccgctttagctcgatcattcttccattgctgcccactgagctggaccttcgcac 360
Db 60 TATACCGCTTGTAGTCGATCATCTTTCACCTTGTGCCACTGAGTCGACCTTCGCAC 1

RESULT 8

US-10-040-862-248/c
; Sequence 248, Application US/10040862
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 248
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(453)
; OTHER INFORMATION: n = A,T,C or G
US-10-040-862-248

Query Match 69.4%; Score 323.4; DB 7; Length 453;
Best Local Similarity 99.1%; Pred. No. 1.1e-88;
Matches 346; Conservative, 0; Mismatches 1; Indels 2; Gaps 2;
Qy 1 atcaagaataagatgctgggcaataatacttctacccatcccccatacttta 60

Db 347 ATCAAGAAACATAGAGTT-GGGCAATATACCTTCACTCCATCCACCCAAATCTTA 289
Qy 61 cttactcatctctcatttaatttgggaaatcatcagaagatgtgtctgttaagta 120
Db 288 CTCTACTCACTCTCACTCTCACTTAATTTTGGGAATCATCAGAGATGTGTCTGTGAGTA 229
Qy 121 agagattaaagaataagctttttgacccttgcacacaccccatgcccagggtgacac 180
Db 238 AGAGATTAAAGAAATAAGCTTTTGTACCCCTGCCAAACACCCCATGCCAGGGTGGTCAC 169
Qy 181 cctccatacaataacatccaggagagtaagtgtgcccctttctgtatgccgttaattctgcc 240
Db 168 CTCTCAATACATAACATCCAGGAAGAGTAAGTTGCCCTTTCTGATGCCGTAATCTGCC 109
Qy 241 atcatcttccatcttccagctctcttccattgctgcccactgagctggaccttcgcac 300
Db 108 ATCATCTTCCCATCTTCCAGTCT-CTTCCATGCGAAGTCACAATCTGGGTCTCAGGAT 50
Qy 301 tataccgctttagctcgatcattcttccattgctgcccactgagctggaccttcgcac 349
Db 49 TATACCGCTTGTAGTCGATCATCTTTCACCTTGTGCCACTGAGCTG 1

RESULT 9

US-10-040-862-5273/c
; Sequence 5273, Application US/10040862
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5273
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (420)

OTHER INFORMATION: n=A,T,C or G
US-10-040-862-5273

Query Match 69.4%; Score 323.4; DB 7; Length 453;
Best Local Similarity 99.1%; Pred. No. 1.1e-88;
Matches 346; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 1 atcaagaacacatagattcgggcaataatacattcattcctaccaccccaaatctta 60
Db 347 ATCAAGAACAATAGAGTT-GGGCAATATACITCATCTACCCATCCCAACCAATCTTA 289
Qy 61 ctctactcattctctcattcatttgggaaatcatcagaagatgtgttcgttgagta 120
Db 288 CTCTACTCATCTCATCTCATTCAATTTTGGAAATCATCAGAAGATGTGTTCTTGAGTA 229
Qy 121 agagattaaagaataaagctttttgacccctgcacaccccccatccagaggtggtcac 180
Db 228 AGAGATTAAAGAATAAGCTTTTTCACCCCTGCCAACACCCCATGCCAGGGTGGTAC 169
Qy 181 cctcaatacaataacatgcaggaagatgaattgtccctttctgtatgcgtaactgcc 240
Db 168 CCTCANTACAAATAGATGCCAGGAGATAGTTCCTTTCTGATGCCGTAATCTGCC 109
Qy 241 atcattctccattctccattctccattctccattctccattctccattctccattctcc 300
Db 108 ATCATCTTCCCATCTTCCAGTCT-CTTCCATTGCCAAGTCACAACTCTGGGCTCAGGAT 50
Qy 301 tataccgtcttagctcgtatcattgtcttccattgtgacctgagctg 349
Db 49 TATACCCGTCTTAGCTTCGATCATCTGCTTTCACCTTGTGCCACTGAGCTG 1

RESULT 10
US-10-137-036-34/C
Sequence 34, Application US/10137036
GENERAL INFORMATION:

APPLICANT: Petera, Ranjan
APPLICANT: Rice, Stephen
APPLICANT: Eagleton, Claire
APPLICANT: Lasham, Annette
APPLICANT: Wood, Marion
APPLICANT: Visser, Elizabeth
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Modification of Gene Expression
FILE REFERENCE: 11000.1036c4
CURRENT APPLICATION NUMBER: US/10/137,036
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: PCT/NZ 01/00115
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: U.S. No. 09/724,624
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: U.S. No. 09/598,401
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: PCT/NZ00/00018
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: U.S. No. 60/146,591
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: U.S. No. 09/276,599
PRIOR FILING DATE: 1999-03-25
NUMBER OF SEQ ID NOS: 143
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 34
LENGTH: 1274
TYPE: DNA
ORGANISM: Eucalyptus grandis
US-10-137-036-34

Query Match 9.8%; Score 45.8; DB 7; Length 1274;
Best Local Similarity 57.8%; Pred. No. 0.0011;
Matches 100; Conservative 0; Mismatches 72; Indels 1; Gaps 1;

Qy 200 ccaggaagagtaagtgtcccttctctgatgcgtaattcgtccatcatcttcccatctcca 259
Db 1241 CCAAGTGGAGGTGGACTCTCTTGAATTTATAATCGGCAAGGTTGGCCATCTTCCA 1182
Qy 260 gtctcttccattgcaagtcacaatctgggtctcagggaattatacccttcttagtctcg 319
Db 1181 G-CTGCTTGGCAGCAAGATAAGCCTCTGCTGGTCCGAGGATCCCTTCTTGTCTCTGG 1123
Qy 320 atcattgtcttctcattgtgcacatgagctgcaccttcgacctgggagggagt 372
Db 1122 ATTTTGTCTTCAATATCAATCTGTCGGAGCTTTCCACCTCGAGGGTGTAT 1070

RESULT 11
US-09-620-393B-2769/C
Sequence 2769, Application US/09620393B
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nickolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
TITLE OF INVENTION: THEREBY
FILE REFERENCE: 2750-1068P
CURRENT APPLICATION NUMBER: US/09/620,393B
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 9948
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2769
LENGTH: 442
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURES:
NAME/KEY: misc.feature
LOCATION: 1..442
OTHER INFORMATION: any n = a, g, c, t, unknown, or other
FEATURES:
NAME/KEY: misc.feature
LOCATION: 1..442
OTHER INFORMATION: Ceres Seq. ID 1384117
US-09-620-393B-2769

Query Match 9.5%; Score 44.2; DB 5; Length 442;
Best Local Similarity 57.2%; Pred. No. 0.0023;
Matches 99; Conservative 0; Mismatches 73; Indels 1; Gaps 1;
Qy 200 ccaggaagagtaagtgtcccttctctgatgcgtaattcgtccatcatcttcccatctcca 259
Db 274 CCAAGTGGAGGTGGACTCTCTTGAATTTATAATCGGCAAGGTTGGCCATCTTCCA 215
Qy 260 gtctcttccattgcaagtcacaatctgggtctcagggaattatacccttcttagtctcg 319
Db 214 G-CTGCTTGGCAGCAAGATAAGCCTCTGCTGGTCCGAGGATCCCTTCTTGTCTCTGG 156
Qy 320 atcattgtcttctcattgtgcacatgagctgcaccttcgacctgggagggagt 372
Db 155 ATCTTGGCCTTAACGTGTGCGATGTGTCGGAGCTTTCCACCTCGAGGGTGTAT 103

RESULT 12
US-09-620-393B-4443/C
Sequence 4443, Application US/09620393B
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nickolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
TITLE OF INVENTION: THEREBY
FILE REFERENCE: 2750-1068P
CURRENT APPLICATION NUMBER: US/09/620,393B
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 9948
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4443
LENGTH: 501
TYPE: DNA
ORGANISM: Arabidopsis thaliana

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..501
OTHER INFORMATION: any n = a, g, c, t, unknown, or other
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..501
OTHER INFORMATION: Ceres Seq. ID 1388198
US-09-620-393B-4443

Query Match          9.5%; Score 44.2; DB 5; Length 501;
Best Local Similarity 57.2%; Pred. No. 0.0024;
Matches 99; Conservative .0; Mismatches 73; Indels 1; Gaps 1;

QY    200   ccaggaagaagttaaggcttgcctttcgtatgacggtaattgcccatcatcttccccattccca 259
      ||||| | ||||| | ||||| ||||| | ||||| ||||| | ||||| ||||| |
DB     274   CCMAAGTGGAGGGTGATTCCTCTCGTGNATTGTAAATCACCACCAACGTCACGGCCATCCTCTA 215
      ||||| | ||||| | ||||| ||||| | ||||| ||||| | ||||| ||||| |
      260   atatatcgatccgatccagatcttcagatcttcagaggaattacccttatctatctcg 319

```

```

Ddb      214 G-CTGCTTCGGCGGAAAAATAAGCCCTTCTGCTATCCGGAGGAATGCCCTTCTTATCCTGG   136
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy       320 atcattgcttttcacttgtgcacctgagctggaccttcgcacctgtggaggaggt    372
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       155 ATCTTGCCCTTAACGTGTGTGATGGTGTGGAGCTTCCACCTCGAGGGTGAT    103

RESULT 13
US-09-975-254-22419/c
; Sequence 22419, Application US/09975254
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

```

```

; CURRENT APPLICATION NUMBER: US/09/975,254
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/09/263,191
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 31255
; SEQ ID NO 22419
; LENGTH: 243
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700960992H1
US-09-975-254-22419

Query Match          9.3%   Score 43.4   DB 5:   Length 243:
              57.4%   Prod No 0.0032

```

[illegible]

Db 94 TCAATTGTATCAGAACTTTTCCAAAGCTGATGGTCTTTCCTGTCA 46

RESULT 14

US-09-620-393B-1362/c

: Sequence 1362, Application US/09620393B

: GENERAL INFORMATION:

[illegible]

Search completed: July 29, 2002, 19:27:19
Job time: 7610 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 16:37:39 ; Search time 4748.22 seconds
(without alignments)
2123.293 Million cell updates/sec

Title: US-09-880-107-2492

Perfect score: 466

Sequence: 1 atcaagaacacatagattc.....gggtaagggtggtggttt 466

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US100_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US101_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6023_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	465	100.0	466	33	US-09-880-107-2492
2	430.2	85.9	7315	63	US-60-248-505-445
3	430.2	85.9	9250	63	US-60-248-505-278
4	430.2	85.9	125235	63	US-60-245-225-89
5	399.8	85.8	833	1	PCT-US00-05882-475
6	399.8	85.8	833	34	US-09-925-301-475
7	398.6	85.5	798	37	US-10-044-090-482
8	398.6	85.5	798	61	US-60-222-113-231
9	398.6	85.5	798	65	US-60-260-483-482
10	398.6	85.5	951	56	US-60-172-373-12767
11	398.6	85.5	2697	59	US-60-209-009-485
12	398.6	85.5	2697	59	US-60-209-106-423
13	398.6	85.5	2697	60	US-60-213-360-7551
14	398.6	85.5	2697	66	US-60-278-258-16025
15	397	85.2	777	13	US-08-978-289-1
16	397	85.2	777	36	US-09-994-444-1
17	397	85.2	806	23	US-09-606-680-4218
18	397	85.2	32659	59	US-60-206-020-98
19	350.2	83.7	797	18	US-09-471-275-1379
20	379	81.3	588	37	US-10-029-386-10199
21	379	81.3	588	37	US-10-029-386-12723
22	334.4	71.8	474	30	US-09-796-692-205
23	334.4	71.8	474	30	US-09-796-692-646
24	334.4	71.8	474	30	US-09-796-692-4647
25	339.4	70.7	568	25	US-09-644-870-6673
26	325	69.7	384	18	US-09-431-517-10267
27	323.4	69.4	453	30	US-09-796-692-248
28	323.4	69.4	453	30	US-09-796-692-5273
29	316.6	67.9	583	55	US-60-164-378-219
30	316.2	67.9	371	15	US-09-179-473-3905
31	316.2	67.9	371	17	US-09-328-351-3905

ALIGNMENTS

	Query Match	100.0%	Score 466;	DB 33;	Length 466;
	Best Local Similarity	100.0%;	Pred. No. 7e-131;		
Matches	Conservative	0;	Mismatches	0;	Gaps
Oy	1	atcaagaacaatagagtgcggcaataacttcattccaccctcccacccccaaacttca	60		
Db	1	atcaagaacaatagagtgcggcaataacttcattccaccctcccacccccaaacttca	60		
Oy	61	ctctactcattctcatctaattttggaatacatcagaagatgttgtcttaagta	120		
Db	61	ctctactcattctcatctaattttggaatacatcagaagatgttgtcttgagta	120		
Oy	121	agagattaaaagaataaagctttttgacctctgccaacaccccatgccagggtgggtcac	180		
Db	121	agagattaaaagaataaagctttttgacctctgccaacaccccatgccagggtgggtcac	180		
Oy	181	cctccaatacaataaactgccaggagaagtagtgtgcctttcttgatgccgtaaatctgcc	240		
Db	181	cctccaatacaataaactgccaggagaagtagtgtgcctttcttgatgccgtaaatctgcc	240		
Oy	241	atcattcttcccattcttcagttctcctttccattggaagtccaactctgggtctcaggcat	300		
Db	241	atcattcttcccattcttcagttctcctttccattggaagtccaactctgggtctcaggcat	300		
Oy	301	tatacccgctttagctcgatcattgtctttcacttgtgcacttgagctggagaccttcgac	360		
Db	301	tatacccgctttagctcgatcattgtctttcacttgtgcacttgagctggagaccttcgac	360		

```

Query Match      85.9%; Score 400.2; DB 63; Length 7315;
Best local Similarity 98.5%; Pred. No. 2.6e-110;
Matches 446; Conservative 0; Mismatches 3; Indels 4; Gaps 4

```

Qy	1	atcaagaacacatagagctcgccgcaataactctacccatcccaaccaatttta	60
Db	6855	atcaagaacacatagagct-ggcgcaataactctacccatcccaaccaatttta	6913
Qy	61	ctctactcatctcattctcatttaattttgggaaatcatcagaagatgtgttcgttgagta	120
Db	6914	ctctactcatctcattctcatttaattttgggaaatcatcagaagatgtgttcgttgagta	6973
Qy	121	agagattaaaagaataagctttttgacccctgcacacccccatgccagggtgggtcac	180
Db	6974	agagattaaaagaataagctttttgacccctgcacacccccatgccagggtgggtcac	7033
Qy	181	ctctcaatatcaataacatgcaggaagaagtaagtgtccctttctgatgcgctaactctgcc	240
Db	7034	ctctcaatatcaataacatgcaggaagaagtaagtgtccctttctgatgcgctaactctgcc	7093
Qy	241	atcatctcccatcttcagctctccttctccattgcaagtcaaatctgggtctcagggat	300
Db	7094	atcatctcccatcttcagctct-ctttccattgcaagtcaaatctgggtctcagggat	7152
Qy	301	tatacccgcttagtctcgataattgtttcaatttgcacatgacatgagcttcgcac	360
Db	7153	tatacccgcttagtctcgataattgtttcaatttgcacatgacatgagcttcgcac	7212
Qy	361	ctggagaggaggctccttttccctcatcaactgactccacaagaacaaggcgactctct	420
Db	7213	ct-ggagaggaggctccttttccctcatcaactgactccacaagaacaaggcgactctct	7271
Qy	421	catcatggggtcttcaccatcttcagggttaag	453
Db	7272	catcatggggtcttcaccatcttcagggttaag	7303

RESULT 3
US-60-248-505-278
: Sequence 278, Application US/60248505
: GENERAL INFORMATION:
: APPLICANT: Beasley, Ellen
: TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED

; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: C1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 278
; LENGTH: 9250
; TYPE: DNA
; ORGANISM: human
US-60-248-505-278

Query Match 85.9%; Score 400.2; DB 63; Length 9250;
Best Local Similarity 98.5%; Pred. No. 2.9e-110;
Matches 446; Conservative 0; Mismatches 3; Indels 4; Gaps 4;
QY 1 atcaagaacatagagtcgggcaataatacttcatctaccatccaccacaaatctta 60
DB 8790 atcaagaacatagagtt-gggcaataatacttcatctaccatccaccacaaatctta 8848
QY 61 ctctactcatcttcatttaattttgggaaatcatcagaagaatgtgttcgttgagta 120
DB 8849 ctctactcatcttcatttaattttgggaaatcatcagaagaatgtgttcgttgagta 8908
QY 121 agagattaaaagaataagctttttgacccctgccaaaccccatgccagggtgtgtcac 180
DB 8909 agagattaaaagaataagctttttgacccctgccaaaccccatgccagggtgtgtcac 8968
QY 181 ctccaatacaataacatgccaggagtagtggcctttcttgatgcogtaactcgcc 240
DB 8969 ctccaatacaataacatgccaggagtagtggcctttcttgatgcogtaactcgcc 9028
QY 241 atcatcttcccatcttcagctctctttccattgcaagtccaatctgggtctcaggat 300
DB 9029 atcatcttcccatcttcagctctctttccattgcaagtccaatctgggtctcaggat 9087
QY 301 tatacccgcttagtcgatcattgtttcacttgccactgagctggacccttcgcac 360
DB 9088 tatacccgcttagtcgatcattgtttcacttgccactgagctggacccttcgcac 9147
QY 361 ctggaggagagtgcttttccctcatcactgactccaagaacaaggagcagctcct 420
DB 9148 ct-ggaggagggtgcttttccctcatcactgactccaagaacaaggagcagctcct 9206
QY 421 catcactggggttcaccactttcagggttaag 453
DB 9207 catcact-gggcttcaccactttcagggttaag 9238

RESULT 4
US-60-245-225-89
; Sequence 89, Application US/60245225
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: C1000885
; CURRENT APPLICATION NUMBER: US/60/245,225
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 705
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 125235
; TYPE: DNA
; ORGANISM: Human
US-60-245-225-89

Query Match 85.9%; Score 400.2; DB 63; Length 125235;
Best Local Similarity 98.5%; Pred. No. 8.9e-110;

Matches 446; Conservative 0; Mismatches 3; Indels 4; Gaps 4;
QY 1 atcaagaacatagagtcgggcaataatacttcatctaccatccaccacaaatctta 60
DB 95978 atcaagaacatagagtt-gggcaataatacttcatctaccatccaccacaaatctta 96036
QY 61 ctctactcatcttcatttaattttgggaaatcatcagaagaatgtgttcgttgagta 120
DB 96037 ctctactcatcttcatttaattttgggaaatcatcagaagaatgtgttcgttgagta 96096
QY 121 agagattaaaagaataagctttttgacccctgccaaaccccatgccagggtgtgtcac 180
DB 96097 agagattaaaagaataagctttttgacccctgccaaaccccatgccagggtgtgtcac 96156
QY 181 ctccaatacaataacatgccaggagtagtggcctttcttgatgcogtaactcgcc 240
DB 96157 ctccaatacaataacatgccaggagtagtggcctttcttgatgcogtaactcgcc 96216
QY 241 atcatcttcccatcttcagctctctttccattgcaagtccaatctgggtctcaggat 300
DB 96217 atcatcttcccatcttcagctctctttccattgcaagtccaatctgggtctcaggat 96275
QY 301 tatacccgcttagtcgatcattgtttcacttgccactgagctggacccttcgcac 360
DB 96276 tatacccgcttagtcgatcattgtttcacttgccactgagctggacccttcgcac 96335
QY 361 ctggaggagagtgcttttccctcatcactgactccaagaacaaggagcagctcct 420
DB 96336 ct-ggaggagggtgcttttccctcatcactgactccaagaacaaggagcagctcct 96394
QY 421 catcactggggttcaccactttcagggttaag 453
DB 96395 catcact-ggggttcaccactttcagggttaag 96426

RESULT 5
PCT-US00-05882-475/c
; Sequence 475, Application PC/TUS0005882
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
; FILE REFERENCE: PA106PCT
; CURRENT APPLICATION NUMBER: PCT/US00/05882
; CURRENT FILING DATE: 2000-03-08
; EARLIER FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 475
; LENGTH: 833
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (58)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (73)
; OTHER INFORMATION: n equals a,t,g, or c

: PCT-US00-05882-475

Query Match	85.8%	Score	399.8	DB	1	Length	833;
Best Local Similarity	98.2%	Pred. No.	1.4e-110				
Matches	445; Conservative	1;	Mismatches	3;	Indels	4;	Gaps
Qy	1	atcaagaacatagagttcgggcaatactacttccaccatcccccaaatctta	60				
Db	751	ATCAAGAACAATAGAGTT -GGGCAATATACTTCATCCTACCCATCCCACCCAATCTTA	693				
Qy	61	ctctaactcatcttcatttaatttttggaataatcataagagatgttgttcgaqta	120				
Db	692	CTCTACTCATCTCATTTCTTAAATTTGGGAATCATCAGAAGATGTGTCTGCITGAGTA	633				
Qy	121	agagattaaaagaataagctttttgaccttcgcaaccccacgccagggtgtgcac	180				
Db	632	AGAGATTAAAGNAATAGCTTTTGACCCTCGCAACACCCCATGCCAGGTGGTTCAC	573				
Qy	181	cctccaataacaatacattgcaggaaagtagtgcctttctgtatgcagttaactgcac	240				
Db	572	CCTCCAATAACAATAASATGCCAGGAAGTAGTAGTGGCCCTTCTGTATGCCGTAAATCGCC	513				
Qy	241	atactattcccatctccagctcctcttccattgcaagtccaatctgggtctcagggat	300				
Db	512	ATCATCTTCCCAATCTCCAGTCT -CTTTCCATTCGAGTCAACAATCTGGGTCTCAGGGAT	454				
Qy	301	tatacccgctctagcttcgatcattgctttcactgtgccactgagctggaacttcgac	360				
Db	453	TATACCCGCTTGTAGTCTCGATCATTTGCTTCACTTGTGCCACTGAGCTGGACCTTCGCAC	394				
Qy	361	ctggggaggagtgctcttttgcctcatcactgtactccacaagaacaaaggcgagctct	420				
Db	393	CT -GGAGAGGTGCCCTCTTTTGCTCATCACCTTGACTCCCAAGAAACAAGGCGACGTCTCT	335				
Qy	421	catcactggggcttcaccactcttcaggggtaag	453				
Db	334	CATCACT -GGGCTTCACCACTTTCAGGGTAAGG	303				

RESULT 6

US-09-925-301-475/c
: Sequence 475, Application US/09925301

```

sequence 4757. Application: 657/000000000
/
GENERAL INFORMATION:
/
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/
/ FILE REFERENCE: PA106
/
/ CURRENT APPLICATION NUMBER: US/09/925,301
/ CURRENT FILING DATE: 2001-08-10
/
/ PRIOR APPLICATION NUMBER: PCT/US00/05882
/ PRIOR FILING DATE: 2000-03-08
/
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/
/ NUMBER OF SEQ ID NOS: 1694
/
/ SOFTWARE: Patentin ver. 2.0
/
/ SEQ ID NO 475
/
/ LENGTH: 833
/
/ TYPE: DNA
/
/ ORGANISM: Homo sapiens
/
/ FEATURE:
/
/ NAME/KEY: misc_feature
/
/ LOCATION: (9)
/
/ OTHER INFORMATION: n equals a,t,g, or c
/
/ NAME/KEY: misc_feature
/
/ LOCATION: (15)
/
/ OTHER INFORMATION: n equals a,t,g, or c
/
/ NAME/KEY: misc_feature
/
/ LOCATION: (29)
/
/ OTHER INFORMATION: n equals a,t,g, or c
/
/ NAME/KEY: misc_feature
/
/ LOCATION: (58)
/
/ OTHER INFORMATION: n equals a,t,g, or c
/
/ NAME/KEY: misc_feature
/
/

```

: LOCATION: (73)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-475

Query Match	85.8%	Score	399.8	DB	34	Length	833
Best Local Similarity	98.2%	Pred. No.	1.4e-110				
Matches	445	Conservative	1	Mismatches	3	Indels	4
Gaps							

Qy	1	atcaagaacatagagttcgggcaataatacttcatactcctaccatccacacccaatctta	60
Db	751	ATCAAGAAACATAGAGTT-GGGCAATATACTTCATCTACCCATCCACCCAAATCTTA	693
Qy	61	ctctactcatctcatctcatttaatttgggaaatactcagaagaatgtgtcgttgagta	120
Db	692	CTCTACTCATCTCAATCTCATTAATTTGGGAAATCATCAGAAGATGTGTTGTTGAGTA	633
Qy	121	agagattaaaagaataagcttttgcacctctgccacaccccatgccagggtgggtcac	180
Db	632	AGAGATTAAAAGAAATAGCTTTTTCACCTCGCCACACCCCATGCCAGGGTGGTCAC	573
Qy	191	ctctcaatacataaactgccaggagaagtaagtgcctctcttgatgccataactccc	240
Db	572	GCTCCAATACAAATAAGTCCAGGAAGTAGTGAAGTGCCCTTCTGTATGCCGTAATCTGCC	513
Qy	241	atcattctcccatcttcagctctccttcccatgtcaagtccaaactctgggtctcaggagat	300
Db	512	ATCATCTTCCCATCTTCCAGTCT-CTTTCATTGCAAGTCACAACTCGSGTCTCAGGGAT	454
Qy	321	tataccgcgttagtctcgatcattgtttccaattgtgccactgagctggaccttcgcac	360
Db	453	TATACCCGCTTAGTCTCGATCAATGCTTTTCATTGTGCCACTGAGCTGGACCTTCGGAC	394
Qy	361	tggaggagaggtgctcttttacctcatcacctgactccacaagaacaaggcgagctcct	420
Db	393	TT-GGAGGAGGTGGCTCTTTTGCCCTCATCACCTGACTCCACAAGAAACAGGGCAGCTCT	335
Qy	421	catcactggggcttcaccacttccaggggtaag	453
Db	324	CATCACT-GGGCTTCACCACTTTCAGGGTAAGG	303

RESULT

US-10-044-390-482/c
. Sequence 482. Application US/10044090

```

: sequence 4827, application 05/10044-090
:
: GENERAL INFORMATION:
:
: APPLICANT: Olga Bandman
: TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
: FILE REFERENCE: PA-0028 US
: CURRENT APPLICATION NUMBER: US/10/044,090
: CURRENT FILING DATE: 2002-01-09
: NUMBER OF SEQ ID NOS: 850
: SOFTWARE: PEXE Program
:
: SEQ ID NO 482
:
: LENGTH: 798
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: misc_feature
:
: OTHER INFORMATION: Incyte ID No: 575983CBI
:
: US-10-044-090-482

```

```

Query Match      85.5%  Score 398.6; DB 37; Length 798;
Best Local Similarity 98.2%; Pred. No. 3.le-110;
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;
y      1 atcaagaacatagagttcggcaatactatctatctaccatccaccccaaatctta 60
      |||
b      769 ATCAAAGAACAATAGAGTT-GGGCAATATCTTCTTACCATCCACCCAAATCTTA 651
      |||
y      61 ctctactcatctcatctctcatctaattttgggaaatcatcagaagatgtgttcgtgaata 120
      |||

```

```
Db 650 CTCTACTCATCTCATCTCAATTAATTTGGGAATCATCAGAAAGATGTGTTGCTTGAGTA 591
Qy 121 agagattaaagaataagctttttgacccctgccaacacccatgccagggtggtcac 180
Db 590 AGAGATTAAAGAAATAGCTTTTTCAGCCCTGCCAACACCCATGCCCAGGGTGTCTAC 531
Qy 181 cttccaataacaatacaccaggaagatgaattggccttttctatgccgaatctgcc 240
Db 530 CTCCAATAACAATAGATGCCAGGAAGATGAAGTTGGCCCTTTCTGATGCCGTAATCTGCC 471
Qy 241 atcatcttcccattctccagtcctcttccattgcaagtccaatctgggtctcaggat 300
Db 470 ATCATCTTCCCATCTCCAGTCT-CTTTCCATTGCAAGTCAACAATCTGGGTCACAGGAT 412
Qy 301 tataccgcttagtcgatcatctgcttccattgtgcccactgagctggacccctcgac 360
Db 411 TATACCCGCTTAGTCTCGATCATCTGCTTTTACATTTGTGCCACTGAGCTGGACCTTCGCAC 352
Qy 361 ctggaggaggtgctcttctgctcatcactgactcaccagaaacaaaggagcactcct 420
Db 351 CT-GGAGGAGGTGCTCTTTTGGCTCTATCACCCTGACTCCACAAGAAACAAAGGCGACTCTCT 293
Qy 421 catcactggggttccaccacttccagggttaag 453
Db 292 CATCACT-GGGCTTCACCACCTTTTCAGGGTAAGG 261

RESULT 8
US-60-222-113-231/c
; Sequence 231, Application US/60222113
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 P
; CURRENT APPLICATION NUMBER: US/60/222,113
; CURRENT FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 231
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 575983CB1
US-60-222-113-231

Query Match 85.5%; Score 398.6; DB 61; Length 798;
Best Local Similarity 98.2%; Pred. No. 3.1e-110;
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

Qy 1 atcaagaacatagagttcgggcaatactctcatctccatccatcccccaccccaaatctta 60
Db 709 ATCAAGAAACATAGAGTT-GGGCAATATATCTTCCATCTACCCATCCCAACCAATCTTA 651
Qy 61 ctctactcatctctcatcttaatttgggaatcatcagaagatgtgtcgttgagta 120
Db 650 CTCTACTCATCTCATCTCAATTAATTTGGGAATCATCAGAAAGATGTGTTGCTTGAGTA 591
Qy 121 agagattaaagaataagctttttgacccctgccaacacccatgccagggtggtcac 180
Db 590 AGAGATTAAAGAAATAGCTTTTTCAGCCCTGCCAACACCCATGCCCAGGGTGTCTAC 531
Qy 181 cttccaataacaatacaccaggaagatgaattggccttttctatgccgaatctgcc 240
Db 530 CTCCAATAACAATAGATGCCAGGAAGATGAAGTTGGCCCTTTCTGATGCCGTAATCTGCC 471
Qy 241 atcatcttcccattctccagtcctcttccattgcaagtccaatctgggtctcaggat 300
Db 470 ATCATCTTCCCATCTCCAGTCT-CTTTCCATTGCAAGTCAACAATCTGGGTCACAGGAT 412
Qy 301 tataccgcttagtcgatcatctgcttccattgtgcccactgagctggacccctcgac 360
Db 411 TATACCCGCTTAGTCTCGATCATCTGCTTTTACATTTGTGCCACTGAGCTGGACCTTCGCAC 352
Qy 361 ctggaggaggtgctcttctgctcatcactgactcaccagaaacaaaggagcactcct 420
Db 351 CT-GGAGGAGGTGCTCTTTTGGCTCTATCACCCTGACTCCACAAGAAACAAAGGCGACTCTCT 293
Qy 421 catcactggggttccaccacttccagggttaag 453
Db 292 CATCACT-GGGCTTCACCACCTTTTCAGGGTAAGG 261

RESULT 8
US-60-222-113-231/c
; Sequence 231, Application US/60222113
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 P
; CURRENT APPLICATION NUMBER: US/60/222,113
; CURRENT FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 231
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 575983CB1
US-60-222-113-231

Query Match 85.5%; Score 398.6; DB 61; Length 798;
Best Local Similarity 98.2%; Pred. No. 3.1e-110;
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

Qy 1 atcaagaacatagagttcgggcaatactctcatctccatccatcccccaccccaaatctta 60
Db 709 ATCAAGAAACATAGAGTT-GGGCAATATATCTTCCATCTACCCATCCCAACCAATCTTA 651
Qy 61 ctctactcatctctcatcttaatttgggaatcatcagaagatgtgtcgttgagta 120
Db 650 CTCTACTCATCTCATCTCAATTAATTTGGGAATCATCAGAAAGATGTGTTGCTTGAGTA 591
Qy 121 agagattaaagaataagctttttgacccctgccaacacccatgccagggtggtcac 180
Db 590 AGAGATTAAAGAAATAGCTTTTTCAGCCCTGCCAACACCCATGCCCAGGGTGTCTAC 531
Qy 181 cttccaataacaatacaccaggaagatgaattggccttttctatgccgaatctgcc 240
Db 530 CTCCAATAACAATAGATGCCAGGAAGATGAAGTTGGCCCTTTCTGATGCCGTAATCTGCC 471
Qy 241 atcatcttcccattctccagtcctcttccattgcaagtccaatctgggtctcaggat 300
Db 470 ATCATCTTCCCATCTCCAGTCT-CTTTCCATTGCAAGTCAACAATCTGGGTCACAGGAT 412
Qy 301 tataccgcttagtcgatcatctgcttccattgtgcccactgagctggacccctcgac 360
```

```
Db 411 TATACCCGCTTAGTCTCGATCATCTTTCATCTTGGCCACTGAGCTGACCTTCGCAC 352
Qy 361 ctggaggaggtgctcttctgctcatcactgactcaccagaaacaaaggagcactcct 420
Db 351 CT-GGAGGAGGTGCTCTTTTGGCTCTATCACCCTGACTCCACAAGAAACAAAGGCGACTCTCT 293
Qy 421 catcactggggttccaccacttccagggttaag 453
Db 292 CATCACT-GGGCTTCACCACCTTTTCAGGGTAAGG 261

RESULT 9
US-60-260-483-482/c
; Sequence 482, Application US/60260483
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 P
; CURRENT APPLICATION NUMBER: US/60/260,483
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 482
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 575983CB1
US-60-260-483-482

Query Match 85.5%; Score 398.6; DB 65; Length 798;
Best Local Similarity 98.2%; Pred. No. 3.1e-110;
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

Qy 1 atcaagaacatagagttcgggcaatactctcatctccatccatcccccaccccaaatctta 60
Db 709 ATCAAGAAACATAGAGTT-GGGCAATATATCTTCCATCTACCCATCCCAACCAATCTTA 651
Qy 61 ctctactcatctctcatcttaatttgggaatcatcagaagatgtgtcgttgagta 120
Db 650 CTCTACTCATCTCATCTCAATTAATTTGGGAATCATCAGAAAGATGTGTTGCTTGAGTA 591
Qy 121 agagattaaagaataagctttttgacccctgccaacacccatgccagggtggtcac 180
Db 590 AGAGATTAAAGAAATAGCTTTTTCAGCCCTGCCAACACCCATGCCCAGGGTGTCTAC 531
Qy 181 cttccaataacaatacaccaggaagatgaattggccttttctatgccgaatctgcc 240
Db 530 CTCCAATAACAATAGATGCCAGGAAGATGAAGTTGGCCCTTTCTGATGCCGTAATCTGCC 471
Qy 241 atcatcttcccattctccagtcctcttccattgcaagtccaatctgggtctcaggat 300
Db 470 ATCATCTTCCCATCTTCCAGTCT-CTTTCCATTGCAAGTCAACAATCTGGGTCACAGGAT 412
Qy 301 tataccgcttagtcgatcatctgcttccattgtgcccactgagctggacccctcgac 360
Db 411 TATACCCGCTTAGTCTCGATCATCTGCTTTTACATTTGTGCCACTGAGCTGGACCTTCGCAC 352
Qy 361 ctggaggaggtgctcttctgctcatcactgactcaccagaaacaaaggagcactcct 420
Db 351 CT-GGAGGAGGTGCTCTTTTGGCTCTATCACCCTGACTCCACAAGAAACAAAGGCGACTCTCT 293
Qy 421 catcactggggttccaccacttccagggttaag 453
Db 292 CATCACT-GGGCTTCACCACCTTTTCAGGGTAAGG 261

RESULT 10
US-60-172-373-12767/c
; Sequence 12767, Application US/60172373
```

```

RESULT 11
US-60-209-009-485/c
; Sequence 485, Application US/60209009
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027 P
; CURRENT APPLICATION NUMBER: US/60/209,009
; CURRENT FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 485
; LENGTH: 2697
; TYPE: DNA

```

```

RESULT 12
US-60-269-106-423/c
: Sequence 423, Application US/60209106
: GENERAL INFORMATION:
: APPLICANT: Kaser, Matthew R.
: APPLICANT: Bandman, Olga
: TITLE OF INVENTION: GENES EXPRESSED IN TNU
: TITLE OF INVENTION: MUSCLE CELLS
: FILE REFERENCE: PA-0031 P
: CURRENT APPLICATION NUMBER: US/60/209,106
: CURRENT FILING DATE: 2000-06-02
: NUMBER OF SEQ ID NOS: 499
: SOFTWARE: PERL Program
: SEQ ID NO 423
: LENGTH: 2697
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 208328.
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 1547-1579
: OTHER INFORMATION: a, t, c, g, or other
US-60-269-106-423

```

Query Match 85.5%; Score 398.6; DB 59; Length 2697;
Best Local Similarity 98.2%; Pred. No. 5.2e-110;
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

Qy 1 atcaagaacataaggttcgggcaataatacttccatccaccaccaccccaaatctta 60
Db ATCAAGAACAATAGAGTT-GGGCAATATCTTCATCTACCCATCCACCCCAATCTTA 651

Qy 61 ctctactcatctcatttcatttaattttgggaaatcatcagaagatgtgttcgttgagta 120
Db CTCTACTCATCTCAITCTCAITTAATTTTGGGAAATCATCAGAGATGTGTTCGTGAGTA 591

Qy 121 agagattaaagaataagctttttgacccctgccaacaccccatgccaggggtgcac 180
Db AGAGATTAAAGAATAAGCTTTTGGACCCCTGCCAACACCCCATGCCAGGGTGGTCAC 531

Qy 181 ctctactcatctcatttcatttaattttgggaaatcatcagaagatgtgttcgttgagta 240
Db CTCTACTCATCTCAITCTCAITTAATTTTGGGAAATCATCAGAAGATGTGTTCGTGAGTA 591

Qy 121 agagattaaagaataagctttttgacccctgccaacaccccatgccaggggtgcac 180
Db AGAGATTAAAGAATAAGCTTTTGGACCCCTGCCAACACCCCATGCCAGGGTGGTCAC 531

Qy 181 ctctactcatctcatttcatttaattttgggaaatcatcagaagatgtgttcgttcgac 240
Db CTCTACTCATCTCAITCTCAITTAATTTTGGGAAATCATCAGAAGATGTGTTCGTGAGTA 591

Qy 241 atcatcttcccatcttccagttctcttccatgcaagtcacaatctgggtctcaggat 300
Db ATCATCTTCCCATCTTCCAGTCT-CTTTCATTCGACATCACTGACTCCACAAAGGAGGAGTCTCT 412

Qy 301 tataccctgttagtcgatcatcttccatgcaagtcacaatctgggtctcaggat 360
Db TATACCCGCTTTAGTCTCGATCATCTTTCATCTTTCGCTCATCACTGACTCCACAAAGGAGGAGTCTCT 412

Qy 361 ctggggagggtgctcttctgctcacaactgacccgagctggagggcagctcct 420
Db CTGGGAGGAGGTGCTCTTTCGCTCATCACTGACTCCACAAAGGAGGAGTCTCT 412

Qy 421 catcactggggttcaccactttcagggttaag 453
Db CATCACT-GGGCTTCACCACTTTCAGGGTAAGG 261

RESULT 13
US-60-213-360-7551/c
: Sequence 7551, Application US/60213360
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lal, Preeti
: APPLICANT: diep, Dinh
: TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
: TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymor
: FILE REFERENCE: GX-0014 P
: CURRENT APPLICATION NUMBER: 2000-06-21
: NUMBER OF SEQ ID NOS: 8347
: SOFTWARE: PERL Program
: SEQ ID NO 7551
: LENGTH: 2697
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 208328.1
: NAME/KEY: unsure
: LOCATION: 1547-1579
: OTHER INFORMATION: a, t, c, g, or other
US-60-213-360-7551

Query Match 85.5%; Score 398.6; DB 60; Length 2697;
Best Local Similarity 98.2%; Pred. No. 5.2e-110;
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

Qy 1 atcaagaacataaggttcgggcaataatacttccatccaccaccaccccaaatctta 60
Db ATCAAGAACAATAGAGTT-GGGCAATATCTTCATCTACCCATCCACCCCAATCTTA 651

Qy 61 ctctactcatctcatttcatttaattttgggaaatcatcagaagatgtgttcgttgagta 120
Db CTCTACTCATCTCAITCTCAITTAATTTTGGGAAATCATCAGAGATGTGTTCGTGAGTA 591

Qy 121 agagattaaagaataagctttttgacccctgccaacaccccatgccaggggtgcac 180
Db AGAGATTAAAGAATAAGCTTTTGGACCCCTGCCAACACCCCATGCCAGGGTGGTCAC 531

Qy 181 ctctactcatctcatttcatttaattttgggaaatcatcagaagatgtgttcgttcgac 240
Db CTCTACTCATCTCAITCTCAITTAATTTTGGGAAATCATCAGAAGATGTGTTCGTGAGTA 591

Qy 241 atcatcttcccatcttccagttctcttccatgcaagtcacaatctgggtctcaggat 300
Db ATCATCTTCCCATCTTCCAGTCT-CTTTCATTCGACATCACTGACTCCACAAAGGAGGAGTCTCT 412

Qy 301 tataccctgttagtcgatcatcttccatgcaagtcacaatctgggtctcaggat 360
Db TATACCCGCTTTAGTCTCGATCATCTTTCATCTTTCGCTCATCACTGACTCCACAAAGGAGGAGTCTCT 412

Qy 361 ctggggagggtgctcttctgctcacaactgacccgagctggagggcagctcct 420
Db CTGGGAGGAGGTGCTCTTTCGCTCATCACTGACTCCACAAAGGAGGAGTCTCT 412

Qy 421 catcactggggttcaccactttcagggttaag 453
Db CATCACT-GGGCTTCACCACTTTCAGGGTAAGG 261

RESULT 14
US-60-279-258-16025/c
: Sequence 16025, Application US/60278258
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lal, Preeti
: APPLICANT: diep, Dinh
: TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
: TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
: FILE REFERENCE: GX-0010-1 P
: CURRENT APPLICATION NUMBER: US/60/278,258
: CURRENT FILING DATE: 2001-03-23
: NUMBER OF SEQ ID NOS: 17730
: SOFTWARE: PERL Program
: SEQ ID NO 16025
: LENGTH: 2697
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 208328.1
: NAME/KEY: unsure
: LOCATION: 1547-1579
: OTHER INFORMATION: a, t, c, g, or other
US-60-278-258-16025

Query Match 85.5%; Score 398.6; DB 66; Length 2697;
Best Local Similarity 98.2%; Pred. No. 5.2e-110;
Matches 445; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

Qy 1 atcaagaacataaggttcgggcaataatacttccatccaccaccaccccaaatctta 60
Db ATCAAGAACAATAGAGTT-GGGCAATATATCTTCATCTACCCATCCACCCCAATCTTA 651

Qy 61 ctctactcatctcatttcatttaattttgggaaatcatcagaagatgtgttcgttgagta 120
Db CTCTACTCATCTCAITCTCAITTAATTTTGGGAAATCATCAGAAGATGTGTTCGTGAGTA 591

Qy 121 agagattaaagaataagctttttgacccctgccaacaccccatgccaggggtgcac 180
Db AGAGATTAAAGAATAAGCTTTTGGACCCCTGCCAACACCCCATGCCAGGGTGGTCAC 531

Qy 181 cctccataataacatccaggaagagtaagtgcctttctgagatcccaataatctgcc 240
Db 530 CTTCAATAATAAGATGCCAGGAAGAGTAAGTTCCTTCTGATGCCGTAATCTGCC 471
Qy 241 atcatcttccatcttccagctctcttccattgcaagtccaaatctgggtctcagggat 300
Db 470 ATCACTTCCCATCTTCAGTCT -CTTTCATTGCAAGTCACAATCTGGTCTCAGGAT 412
Qy 301 tataccgctttagtctgatcattgcttctactctggtccactgagctggaccttcgac 360
Db 411 TATACCGCTTCTAGTCTGATCATCTTTCACCTTGTGACCTGAGCTGGACCTTCGCAC 352
Qy 361 ctggagagagctcctcttctgctcactcactgactccacaaagaaacaaggagcagctcct 420
Db 351 CT-GGAGGAGGTGCTCTTTCCTCTCATCTGACTTCCACAAAGAAACAGGCGCTCT 293
Qy 421 catcactggggtctcaccacttccagggttaag 453
Db 292 CATCACT-GGGCTTCACCACTTTTCAGGGTAAGG 261

RESULT 15

US-08-978-289-1/c
; Sequence 1, Application US/08978289
; GENERAL INFORMATION:
; APPLICANT: Bates, Elizabeth E.M.
; APPLICANT: de Saint-Vis, Blandine M.
; APPLICANT: Caux, Christophe
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Bancheureau, Jacques
; TITLE OF INVENTION: ISOLATED MAMMALIAN DENDRITIC CELL GENES:
; TITLE OF INVENTION: RELATED REAGENTS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,289
; FILING DATE: 25-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/032,767
; FILING DATE: 11-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,806
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0669K1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)952-9196
; TELEFAX: (650)952-1204
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 777 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 19..513

US-08-978-289-1

Query Match 85.2%; Score 397; DB 13; Length 777;
Best Local Similarity 98.0%; Pred. NO. 9.4e-110;
Matches 444; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
Qy 1 atcaagaacacatagagttcgggcaataataactctacccatccaccaccccaatctta 60
Db 591 ATCAAGAAACATAGAGTT-GGGCAATATACTTCACTTACCCATCCCAACCAATCTTA 633
Qy 61 cttactcattctcattctcatttaatttgggaaatcatcagaagatgtgttgagta 120
Db 632 CTCTACTCATCTCATCTCTCAATTAATTTGGGAAATCATCAGAAGATGTGTCTGAGTA 573
Qy 121 agagattaaagaataaagctttttgacccctgcccacaccccatgccaggggtggtcac 180
Db 572 AGAGATTAAAGAAATAAGCTTTTTCACCCCTGCCAACACCCCATCCCGAGGTGTCAC 513
Qy 181 cctccataataacatccaggaagagtaagtgcctttctgagatcccaataatctgcc 240
Db 512 CTTCAATAATAAGATGCCAGGAAGAGTAAGTTCCTTCTGATGCCGTAATCTGCC 453
Qy 241 atcatcttccatcttccagctctcttccattgcaagtccaaatctgggtctcaggat 300
Db 452 ATCACTTCCCATCTTCAGTCT -CTTTCATTGCAAGTCACAATCTGGTCTCAGGAT 394
Qy 301 tataccgctttagtctgatcattgcttctactctggtccactgagctggaccttcgac 360
Db 393 TATACCGCTTCTAGTCTGATCATCTTTCACCTTGTGACCTGAGCTGGACCTTCGCAC 334
Qy 351 ctggagagagctcctcttctgctcactcactgactccacaaagaaacaaggagcagctcct 420
Db 333 CT-GGAGGAGGTGCTCTTTCCTCTCATCTGACTTCCACAAAGAAACAGGCGCTCT 275
Qy 421 catcactggggtctcaccacttccagggttaag 453
Db 274 CATCACT-GGGCTTCACCACTTTTCAGGGTAAGG 243

Search completed: July 29, 2002, 19:20:56
Job time: 9797 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2002, 16:29:39 ; Search time 100.77 Seconds
(without alignments)
1135.906 Million cell updates/sec

Title: US-09-880-107-2492
Perfect score: 466
Sequence: 1 atcaagaacatagattc.....gggtaagggtggatgtctt 466

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents.NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43.4	9.3	7218	1	US-08-232-463-14
2	42.2	9.1	3840	1	US-08-462-092-1
3	42.2	9.1	3840	3	US-08-746-822-1
4	42.2	9.1	3840	3	US-09-094-350-1
5	42.2	9.1	3840	6	5510474-1
6	37.4	8.0	508	3	US-09-051-969A-1
7	36.6	7.9	2127	2	US-08-505-486-95
8	36.6	7.9	2127	3	US-08-801-028-95
9	36.6	7.9	2127	3	US-09-340-154-95
10	36.6	7.9	2127	5	PCT-US95-09338-95
11	36.6	7.9	2127	5	PCT-US95-09339-95
12	35.2	7.6	471	2	US-09-070-060-9
13	35.2	7.6	471	3	US-09-357-746-9
14	35	7.5	1228	3	US-08-505-486-92
15	35	7.5	1228	3	US-08-801-028-92
16	35	7.5	1228	3	US-09-340-154-92
17	35	7.5	1228	5	PCT-US95-09338-92
18	35	7.5	1228	5	PCT-US95-09339-92
19	34.6	7.4	581	2	US-08-834-306-22
20	34.6	7.4	581	4	US-08-993-674A-22
21	34.6	7.4	581	4	US-08-505-486-93
22	34.2	7.3	1154	3	US-08-801-028-93
23	34.2	7.3	1154	3	US-09-340-154-93
24	34.2	7.3	1154	5	PCT-US95-09338-93
25	34.2	7.3	1154	5	PCT-US95-09339-93
26	34.2	7.3	1154	2	US-08-505-486-96
27	34.2	7.3	2022	2	

C 28	34.2	7.3	2022	3	US-08-801-028-96	Sequence 96, Appl
C 29	34.2	7.3	2022	3	US-09-340-154-96	Sequence 96, Appl
C 30	34.2	7.3	2022	5	PCT-US95-09338-96	Sequence 96, Appl
C 31	34.2	7.3	2022	5	PCT-US95-09339-96	Sequence 96, Appl
C 32	33.8	7.3	231	1	US-08-450-834-1	Sequence 1, Appl
C 33	33.8	7.3	831	1	US-08-450-834-5	Sequence 5, Appl
C 34	33.6	7.2	487	3	US-09-051-969A-2	Sequence 2, Appl
C 35	32.4	7.0	4673	1	US-07-638-431-1	Sequence 1, Appl
C 36	32.4	7.0	4673	5	PCT-US92-00018-1	Sequence 22, Appl
C 37	31	6.7	707	3	US-08-726-306A-22	Sequence 6, Appl
C 38	30.8	6.6	418	3	US-09-141-000-6	Sequence 4, Appl
C 39	30.8	6.6	458	3	US-09-141-000-4	Sequence 24, Appl
C 40	30.8	6.6	1501	2	US-08-145-658D-24	Sequence 73, Appl
C 41	30.2	6.5	15766	4	US-09-338-907-73	Sequence 73, Appl
C 42	30.2	6.5	15766	4	US-09-218-207-73	Sequence 183, App
C 43	30.2	6.5	37950	4	US-09-338-907-183	Sequence 183, App
C 44	30.2	6.5	37950	4	US-09-218-207-183	Sequence 30, Appl
C 45	30	6.4	1622	1	US-08-216-276A-30	

ALIGNMENTS

RESULT 1
US-08-232-463-14
: Sequence 14, Application US/08232463
: Patent No. 5670367
: GENERAL INFORMATION:
: APPLICANT: DORNER, F.
: APPLICANT: SCHEIFLINGER, F.
: APPLICANT: FALKNER, F. G.
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232,463
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/935,313
: FILING DATE:
: APPLICATION NUMBER: EP 91 114 300.6
: FILING DATE: 26-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30472/114 IMMU
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: pizgpt-Fls
US-08-232-463-14

[illegible]

```

RESULT      2
US-08--462-092-1/c
; Sequence 1, Application US/08462092
; Patent No. 5614399
;
; GENERAL INFORMATION:
;
; APPLICANT: Peter H. Quail
; APPLICANT: Alan H. Christensen
; APPLICANT: Howard P. Hershey
; APPLICANT: Robert A. Sharrock
; APPLICANT: Thomas D. Sullivan
;
; TITLE OF INVENTION: PLANT UBIQUITIN PROMOTER SYSTEM
;
; NUMBER OF SEQUENCES: 2
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeff Lloyd
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,092
; FILING DATE: 5-JUNE-1995
; CLASSIFICATION: 435
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Lloyd, Jeff
; REGISTRATION NUMBER: 35,589
; REFERENCE/DOCKET NUMBER: 08/462,092
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
;
; INFORMATION FOR SEQ ID NO: 1:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3840 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
;

```

```

: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1993..3591
: US-08-462-092-1

Query Match          9.1%; Score 42.2; DB 1; Length 3840;
Best Local Similarity 55.1%; Pred. No. 0.00017;
Matches 103; Conservative 0; Mismatches 83; Indels 1; Gaps

QY 200 ccaggagagtaagtgccttcttgatgccgtaatactgccaatacttcccatcttcca 259
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3568 CAAGGTGGAGGTGCTCTCCTTTGGATGTTGTAGCTGTCTAGGTGCGGCCATCCTCCA 3509

QY 260 gtctctttccattgcaagtcaaatctgggtctcaggagattataccgcgtcttagtctcg 319
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3568 G-CGTGTTGCCGCGAAGATCAGACGCTGCTGGTCCGGGGGATGCCCTCCTTGTCTCTGG 3450

QY 320 atcattgtcttactgtgccactgagctggaccttcgaacctgggaggaggtgcctctt 379
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3449 ATCTTCGCTTCACATGTCGATGTGTCGAGCTCTCAACCTCCAGGGTTATGGTITTT 3390

QY 380 tgcctca 386
      |||
Db 3389 CCAGTCA 3383

RESULT 3
US-08-746-822-1/c
: Sequence 1, Application US/08746822
: Patent No. 6020190
: GENERAL INFORMATION:
: APPLICANT: Peter H. Quail
: APPLICANT: Alan H. Christensen
: APPLICANT: Howard P. Hershey
: APPLICANT: Robert A. Sharrock
: APPLICANT: Thomas D. Sullivan
: TITLE OF INVENTION: PLANT UB1QUITIN PROMOTER SYSTEM
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Jeff Lloyd
: STREET: 2421 N.W. 41st Street, Suite A-1
: CITY: Gainesville
: STATE: FL
: COUNTRY: USA
: ZIP: 32606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/746,822
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/462,092
: FILING DATE: 5-JUNE-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Lloyd, Jeff
: REGISTRATION NUMBER: 35,589
: REFERENCE/DOCKET NUMBER: 08/462,092
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 904-375-8100
: TELEFAX: 904-372-5800
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3840 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single

```

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1993..3591
US-08-746-822-1

Query Match 9.1% Score 42.2; DB 3; Length 3840;

Best Local Similarity 55.1% Pred. No. 0.00017;
Matches 103; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

Qy 200 ccaggaagagtaagtgcctttctgtagccgtaataatcgccatcatcttcccatctcca 259
Db 3568 CAAGGTGGAGGGTCTCTCTTTTGGATGTTGTAGTCTGCTAGGGTGGCGGCATCTCCA 3509
Qy 260 gtctcttccattgcaatgcaatctggtctcaaggattataccggtcttagctcg 319
Db 3508 G-CGTGTTCCGGCGAAGATCAGACGCTGCTGTCGGGGGGATGCCCTCTTGTCTCG 3450
Qy 320 atcattgcttccattgctccactgagctggaccttcgacccctggagaggtgcctctt 379
Db 3449 ATCTCGCTTCACATTTGCTGATGTTGTCGAGCTCTCAACCTCCAGGGTTATGGTTTT 3390
Qy 380 tgcctca 386
Db 3389 CCAGTCA 3383

RESULT 4

US-09-094-350-1/c
Sequence 1, Application US/09094350

Patent No. 6054574

GENERAL INFORMATION:

APPLICANT: Peter H. Quail

APPLICANT: Alan H. Christensen

APPLICANT: Howard P. Hershey

APPLICANT: Robert A. Sharrock

APPLICANT: Thomas D. Sullivan

TITLE OF INVENTION: PLANT UBIQUITIN PROMOTER SYSTEM

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jeff Lloyd

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION NUMBER: US/09/094,350

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/746,822

FILING DATE:

APPLICATION NUMBER: US/08/462,092

FILING DATE: 5-JUNE-1995

ATTORNEY/AGENT INFORMATION:

NAME: Lloyd, Jeff

REGISTRATION NUMBER: 35,589

REFERENCE/DOCKET NUMBER: 08/462,092

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3840 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1993..3591
US-09-094-350-1

Query Match 9.1% Score 42.2; DB 3; Length 3840;

Best Local Similarity 55.1% Pred. No. 0.00017;
Matches 103; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

Qy 200 ccaggaagagtaagtgcctttctgtagccgtaataatcgccatcatcttcccatctcca 259
Db 3568 CAAGGTGGAGGGTCTCTCTTTTGGATGTTGTAGTCTGCTAGGGTGGCGGCATCTCCA 3509
Qy 260 gtctcttccattgcaatgcaatctggtctcaaggattataccggtcttagctcg 319
Db 3508 G-CGTGTTCCGGCGAAGATCAGACGCTGCTGTCGGGGGGATGCCCTCTTGTCTCG 3450
Qy 320 atcattgcttccattgctccactgagctggaccttcgacccctggagaggtgcctctt 379
Db 3449 ATCTCGCTTCACATTTGCTGATGTTGTCGAGCTCTCAACCTCCAGGGTTATGGTTTT 3390
Qy 380 tgcctca 386
Db 3389 CCAGTCA 3383

RESULT 5

5510474-1/c

Patent No. 5510474

APPLICANT: QUAIL, PETER H.; CHRISTENSEN, ALAN H.; HERSHEY,

HOWARD P.; SHARROCK, ROBERT A.; SULLIVAN, THOMAS D.

TITLE OF INVENTION: PLANT UBIQUITIN PROMOTER SYSTEM

NUMBER OF SEQUENCES: 12

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/296,268

FILING DATE: 25-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 191,134

FILING DATE: 03-FEB-1994

APPLICATION NUMBER: 76,363

FILING DATE: 11-JUN-1993

APPLICATION NUMBER: 670,496

FILING DATE: 15-MAR-1991

APPLICATION NUMBER: 194,824

FILING DATE: 17-MAY-1988

SEQ ID NO:1:

LENGTH: 3840

5510474-1

Query Match 9.1% Score 42.2; DB 6; Length 3840;

Best Local Similarity 55.1% Pred. No. 0.00017;

Matches 103; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

Qy 200 ccaggaagagtaagtgcctttctgtagccgtaataatcgccatcatcttcccatctcca 259
Db 3568 CAAGGTGGAGGGTCTCTCTTTTGGATGTTGTAGTCTGCTAGGGTGGCGGCATCTCCA 3509
Qy 260 gtctcttccattgcaatgcaatctggtctcaaggattataccggtcttagctcg 319
Db 3508 G-CGTGTTCCGGCGAAGATCAGACGCTGCTGTCGGGGGGATGCCCTCTTGTCTCG 3450
Qy 320 atcattgcttccattgctccactgagctggaccttcgacccctggagaggtgcctctt 379
Db 3449 ATCTCGCTTCACATTTGCTGATGTTGTCGAGCTCTCAACCTCCAGGGTTATGGTTTT 3390
Qy 380 tgcctca 386

Db 3389 CCAGTCA 3383

|||

RESULT 6
US-09-051-969A-1/c
; Sequence 1, Application US/09051969A
; Patent No. 6063598
; GENERAL INFORMATION:
; APPLICANT: ENENKEL, BARBARA
; APPLICANT: GANNON, FRANK
; APPLICANT: BERGMANN, KLAUS
; APPLICANT: NOE, WOLFGANG
; TITLE OF INVENTION: INTENSIVE HOMOLOGOUS PROMOTER OBTAINED FROM HAMSTERS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/051,969A
; FILING DATE: 1998-09-30
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FLESHNER, RAZ E.
; REGISTRATION NUMBER: 34,331
; REFERENCE/DOCKET NUMBER: 0652.1690000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 508 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA to mRNA
US-09-051-969A-1

Query Match 8.0%; Score 37.4; DB 3; Length 508;
Best Local Similarity 56.8%; Pred. No. 0.0023;
Matches 88; Conservative 0; Mismatches 66; Indels 1; Gaps 1;

QY 218 ccttctgacgcgttaactgccatcattcccatcttcagtcctcttccattgcaa 277
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 206 CCTTTTGATGTTGTAGTCACAAAGTACGGCATCTCCAGT-TGCTTACCAGCAAAG 148

QY 278 gtcaacaattgggtcgaggattataaccgctcttagtcgatcattgtttcactgt 337
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 147 ATCAGCGCTCTGCTGTCAGGAGGAATTCTTCCTATATCCIGGATCTTGGCCCTTACATTT 88

QY 338 gccactgactgaccttcgcacctggaggagt 372
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 87 TCTATAGTGTCCGAGGGTTCACCTCAGCGTGAT 53

RESULT 7
US-08-505-486-95/c
; Sequence 95, Application US/08505486
; Patent No. 5955573
; GENERAL INFORMATION:
; APPLICANT: Jesse M. Jaynes
; TITLE OF INVENTION: UBQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PRO

TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 Thirteenth Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/505,486
FILING DATE: 21-JUL-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/279,472
FILING DATE: 22-JUL-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: WALKER, BARBARA W.
REGISTRATION NUMBER: 35,400
REFERENCE/DOCKET NUMBER: 2093-117A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 2127
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE STRANDED
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: GENOMIC DNA AND OTHER DNA
US-08-505-486-95

Query Match 7.9%; Score 36.6; DB 2; Length 2127;
Best Local Similarity 50.2%; Pred. No. 0.011;
Matches 116; Conservative 0; Mismatches 114; Indels 1; Gaps 1;

QY 142 tttgaccttgccaacccccatgccagggtggtgacccctccaataaacatgcc 201
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2054 ITTTTCAACTACGACCCACTCTTTTAACAGCGGATCACCTCCACGTAGACGAGGAC 1995

QY 202 aggaagaagtaaagttgcccctttctgagcgtaaatgctgcatcttccccatcttccagt 261
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1954 AAATGAGGGTAGACTCTCTCTGGATGTTGTAATCAGCTAGTAGTACGTCCTCTCCA-A 1936

QY 262 ctctcttccattgcaagtcacaaatctgggtctcaggattaccgctcttagtctgat 321
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1935 TGCTTTTCGGGGAAGATAAGCCCTTTGTGATCCGGGGAATCTCTTATCTCTGAT 1876

QY 322 cattgctttcacctgtgccaactgagctgacacctgcacacctggaggagggt 372
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1875 CTAGCCTTAAAGTTGTCGATTGTATCACAACCTTTCCACCTCTAGGGTGAT 1825

RESULT 8
US-08-801-028-95/c
; Sequence 95, Application US/08801028
; Patent No. 6018102
; GENERAL INFORMATION:
; APPLICANT: JOAN GARBARINO
; APPLICANT: JESSE M. JAYNES
; APPLICANT: WILLIAM BELKNAP
; TITLE OF INVENTION: UBQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PRO
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:

APPLICANT: Jesse M. Jaynes
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 Thirteenth Street N.W.
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
OPERATING SYSTEM: IBM COMPATIBLE
SOFTWARE: WordPerfect 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/340,154
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/505,486
FILING DATE: 21-JUL-1995
APPLICATION NUMBER: U.S. 08/279,472
FILING DATE: 22-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: WALKER, BARBARA W.
REGISTRATION NUMBER: 35,400
REFERENCE/DOCKET NUMBER: 2093-117A
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 2127
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE STRANDED
TOPOLOGY: LINEAR
MOLECULE TYPE: GENOMIC DNA AND OTHER DNA
DESCRIPTION: GENOMIC DNA AND OTHER DNA
US-09-340-154-95
Query Match 7.9%; Score 36.6; DB 3; Length 2127;
Best Local Similarity 50.2%; Pred. No. 0.011;
Matches 116; Conservative 0; Mismatches 114; Indels 1; Gaps 1;
QY 142 ttgtgacccctgccaacacccatgcccaggtggtgcacccatcccaataacaatgcc 201
Db 2054 TTTTCAACCTACGACCCACCTCTTTTAAACAGCGGATCCACCTCCAGTAGAGGAGCACC 1995
QY 202 aggaagagtaagttgcccctttctgacggtggtgcacccatcccaataacaatgcc 261
Db 1994 AAATGGAGGAGTAGACTCTCTTGGATGTTGTAATCAGCTAGAGTAGCTCCGCTCCCA-A 1936
QY 262 ctctttccattgcaagtcacaaatctggttcagggattacccgtcttagtctcgat 321
Db 1935 CTGCTTTCGGCGAGATAGCCCTTTGCTGATCCGGGGAAATTCCTTCTTATCTGGAT 1876
QY 322 cattgtttcaactgtgcaactgagctggacccctgcacccctggaggaggt 372
Db 1875 CTTAGCCTTAACGTTGTCGATGATATCAGAACCTTCCACCTCTAGGGTGAT 1825
RESULT 10
PCT-US95-09338-95/c
: Sequence 95, Application PC/TUS9509338
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
: TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
: TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
: GENERAL INFORMATION:

ADDRESSEE: STEVEN J. HULTQUIST
ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
STREET: 200 PARK DRIVE, SUITE 210
STREET: P.O. BOX 14329
CITY: RESEARCH TRIANGLE PARK
STATE: NORTH CAROLINA
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
OPERATING SYSTEM: MACINTOSH
SOFTWARE: M.S. WORD 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,028
FILING DATE: 19-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/279,472
FILING DATE: JULY 22, 1994
APPLICATION NUMBER: 08/225,476
FILING DATE: 04-20-94
APPLICATION NUMBER: 08/225,476
FILING DATE: 04-08-94
APPLICATION NUMBER: 08/039,620
FILING DATE: 06-04-93
APPLICATION NUMBER: 08/148,491
FILING DATE: 11-08-93
APPLICATION NUMBER: 08/148,889
FILING DATE: 11-08-93
ATTORNEY/AGENT INFORMATION:
NAME: WASSERMAN, FRAN S.
REGISTRATION NUMBER: 34,273
REFERENCE/DOCKET NUMBER: 4013-104
TELEPHONE: (919)990-9531
TELEFAX: (919)990-9532
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 2127
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE STRANDED
TOPOLOGY: LINEAR
MOLECULE TYPE: GENOMIC DNA AND OTHER DNA
DESCRIPTION: GENOMIC DNA AND OTHER DNA
US-08-801-028-95
Query Match 7.9%; Score 36.6; DB 3; Length 2127;
Best Local Similarity 50.2%; Pred. No. 0.011;
Matches 116; Conservative 0; Mismatches 114; Indels 1; Gaps 1;
QY 142 ttgtgacccctgccaacacccatgcccaggtggtgcacccatcccaataacaatgcc 201
Db 2054 TTTTCAACCTACGACCCACCTCTTTTAAACAGCGGATCCACCTCCAGTAGAGGAGCACC 1995
QY 202 aggaagagtaagttgcccctttctgacggtggtgcacccatcccaataacaatgcc 261
Db 1994 AAATGGAGGAGTAGACTCTCTTGGATGTTGTAATCAGCTAGAGTAGCTCCGCTCCCA-A 1936
QY 262 ctctttccattgcaagtcacaaatctggttcagggattacccgtcttagtctcgat 321
Db 1935 CTGCTTTCGGCGAGATAGCCCTTTGCTGATCCGGGGAAATTCCTTCTTATCTGGAT 1876
QY 322 cattgtttcaactgtgcaactgagctggacccctgcacccctggaggaggt 372
Db 1875 CTTAGCCTTAACGTTGTCGATGATATCAGAACCTTCCACCTCTAGGGTGAT 1825
RESULT 9
US-09-340-154-95/c
: Sequence 95, Application US/09340154
: Patent No. 6084156
: GENERAL INFORMATION:

```

/ NUMBER OF SEQUENCES: 98
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WORDPERFECT 5.1+
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/09338
/ FILING DATE: 21-JUL-1994
/ PRIOR APPLICATION DATA: 08/279,472
/ APPLICATION NUMBER: 08/279,472
/ FILING DATE: 22-JUL-1994
/ INFORMATION FOR SEQ ID NO: 95:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2127
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: DOUBLE STRANDED
/ TOPOLOGY: LINEAR
/ MOLECULE TYPE:
/ DESCRIPTION: GENOMIC DNA AND OTHER DNA
PCT-US95-09338-95

Query Match 7.9%; Score 36.6; DB 5; Length 2127;
Best Local Similarity 50.2%; Pred. No. 0.011;
Matches 116; Conservative 0; Mismatches 114; Indels 1; Gaps 1;

QY 142 ttgtgacctgtgcaacacccatgccagggtggtcaccctcccaataacaatgcc 201
Db 2054 TTTTCAACCTAGGACCCACTCTTTTAAACAGCGGATCCACCTCCAGTAGAGGAGCACC 1995
QY 202 aggaagagtagtgccctttctgtatgccgttaattcgtcattcccatcttccagt 261
Db 1994 AATGGAGGGTAGACTCCTTCGTGATGTTGTAATCAGCTAGAGTACGTCCTCCCA-A 1936
QY 262 ctcttcttcattgcaagtcacaatctctgggtctcaggagattatccccctcttagctgat 321
Db 1935 CTGCTTTCCGCGGAGATAAGCCTTTCTGATCGGGGGAATTCCTTCCTTATCCTGGAT 1876
QY 322 cattgttctcactgtgcccactgagctggaccctcgaccctcgaggaggt 372
Db 1875 CTTAGCCTTAACGTTGTCGATTGATCAGAACTTCCACCTCTAGGGTGAT 1825

RESULT 11
PCT-US95-09339-95/c
/ Sequence 95, Application PC/TUS9509339
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
/ TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
/ TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
/ NUMBER OF SEQUENCES: 98
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WORDPERFECT 5.1+
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/09339
/ FILING DATE: 21-JUL-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/279,472
/ FILING DATE: 22-JUL-1994
/ INFORMATION FOR SEQ ID NO: 95:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2127
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: DOUBLE STRANDED
/ TOPOLOGY: LINEAR
/ MOLECULE TYPE:
/ DESCRIPTION: GENOMIC DNA AND OTHER DNA

```

PCT-US95-09339-95

```

Query Match 7.9%; Score 36.6; DB 5; Length 2127;
Best Local Similarity 50.2%; Pred. No. 0.011;
Matches 116; Conservative 0; Mismatches 114; Indels 1; Gaps 1;

QY 142 ttgtgacctgtgcaacacccatgccagggtggtcaccctcccaataacaatgcc 201
Db 2054 TTTTCAACCTAGGACCCACTCTTTTAAACAGCGGATCCACCTCCAGTAGAGGAGCACC 1995
QY 202 aggaagagtagtgccctttctgtatgccgttaattcgtcattcccatcttccagt 261
Db 1994 AATGGAGGGTAGACTCCTTCGTGATGTTGTAATCAGCTAGAGTACGTCCTCCCA-A 1936
QY 262 ctcttcttcattgcaagtcacaatctctgggtctcaggagattatccccctcttagctgat 321
Db 1935 CTGCTTTCCGCGGAGATAAGCCTTTCTGATCGGGGGAATTCCTTCCTTATCCTGGAT 1876
QY 322 cattgttctcactgtgcccactgagctggaccctcgaccctcgaggaggt 372
Db 1875 CTTAGCCTTAACGTTGTCGATTGATCAGAACTTCCACCTCTAGGGTGAT 1825

```

RESULT 12

```

US-09-070-960-5/c
/ Sequence 9, Application US/09070060
/ Patent No. 5976849
/ GENERAL INFORMATION:
/ APPLICANT: Hustad, Carolyn M.
/ APPLICANT: Ghildyal, Namit
/ TITLE OF INVENTION: Human E3 Ubiquitin Protein
/ TITLE OF INVENTION: Ligase
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: ZENECA Pharmaceuticals, Inc.
/ STREET: 1800 Concord Pike
/ CITY: Wilmington
/ STATE: DE
/ COUNTRY: USA
/ ZIP: 19850-5437
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/070,060
/ FILING DATE: 30-APR-1998
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/073,839
/ FILING DATE: 05-FEB-1998
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Higgins, Patrick H
/ REGISTRATION NUMBER: 39,709
/ REFERENCE/DOCKET NUMBER: PHM.70312
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 302.886.4889
/ TELEFAX: 302.886.8221
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 471 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: unknown
/ TOPOLOGY: unknown
/ MOLECULE TYPE: cDNA
US-09-070-360-9

```

```

Query Match 7.6%; Score 35.2; DB 2; Length 471;
Best Local Similarity 55.8%; Pred. No. 0.013;

```

Matches 87; Conservative 0; Mismatches 68; Indels 1; Gaps 1;
Oy 218 ccttttgatgcgtaattcccatcatcttcccatctccatctccatctccatgcaa 277
Db 190 CTTTTGAATATTGATGACAGAAAGTACGTCCATCTTCCAG-CTGCTTGCCAGCAAG 132
Oy 278 gtcaaatctgggtctcaggatattaccctcttagctcgatcattgctttcacttgt 337
Db 131 ATCAGTCTCTGCTGATCAGGAGGAATCTCTTATCTCTGATCTTGGCCCTTACATT 72
Oy 338 gccactgagctgaccttcgcacctcggaggaggtg 373
Db 71 TCTATCGTATCCGAGGGTTCACCTCGAGGGTGATG 36

RESULT 13
US-09-357-746-9/c
Sequence 9, Application US/09357746
Patent No. 6087122
GENERAL INFORMATION:
APPLICANT: ZENECA Limited
TITLE OF INVENTION: HUMAN E3 UBQUITIN PROTEIN LIGASE
FILE REFERENCE: PHM.70312.N1
CURRENT APPLICATION NUMBER: US/09/357,746
EARLIER FILING DATE: 1999-07-21
EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
EARLIER FILING DATE: 1998-02-05
EARLIER APPLICATION NUMBER: US No. 608712209/070,060
EARLIER FILING DATE: 1998-04-30
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 9
LENGTH: 471
TYPE: DNA
ORGANISM: Homo sapiens
US-09-357-746-9

Query Match 7.6%; Score 35.2; DB 3; Length 471;
Best Local Similarity 55.8%; Pred. No. 0.013;
Matches 87; Conservative 0; Mismatches 68; Indels 1; Gaps 1;
Oy 218 ccttttgatgcgtaattcccatcatcttcccatctccatctccatctccatgcaa 277
Db 190 CTTTTGAATATTGATGACAGAAAGTACGTCCATCTTCCAG-CTGCTTGCCAGCAAG 132
Oy 278 gtcaaatctgggtctcaggatattaccctcttagctcgatcattgctttcacttgt 337
Db 131 ATCAGTCTCTGCTGATCAGGAGGAATCTCTTATCTCTGATCTTGGCCCTTACATT 72
Oy 338 gccactgagctgaccttcgcacctcggaggaggtg 373
Db 71 TCTATCGTATCCGAGGGTTCACCTCGAGGGTGATG 36

RESULT 14
US-08-505-486-92/c
Sequence 92, Application US/08505486
Patent No. 5955573
GENERAL INFORMATION:
APPLICANT: Jesse M. Jaynes
TITLE OF INVENTION: UBQUITIN-LYTIC PEPTIDE FUSION GENE
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 Thirteenth Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/505,486
FILING DATE: 21-JUL-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/279,472
FILING DATE: 22-JUL-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: WALKER, BARBARA W.
REGISTRATION NUMBER: 35,400
REFERENCE/DOCKET NUMBER: 2093-1117A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 1228
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE STRANDED
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: GENOMIC DNA AND OTHER NUCLEIC ACID
US-08-505-486-92

Query Match 7.5%; Score 35; DB 2; Length 1228;
Best Local Similarity 49.8%; Pred. No. 0.028;
Matches 115; Conservative 0; Mismatches 115; Indels 1; Gaps 1;
Oy 142 ttttgacccctgccacaccccccaggggtggtgcacccctccatccatccatccatgccc 201
Db 1186 TTTTCAACCTACGACCCACTCTTTAACAGCGGATCCACCACCGAGGAGGACG 1127
Oy 202 aggaagagtaagtgccttcttcgataatcgcacatctccatctccatctccatg 261
Db 1126 AGATGGAGAGTGTACTCTCTTGATGTTGTAGTCGGCAAGAGTACGACCATCTCAAG- 1066
Oy 262 ctctttccattgcaagtcacaatctgggtctcaggagattataccctcttagctcgtat 321
Db 1067 CTGCTTTCCGGCGAAATCAAAACGCTGCTGCTGGGGAATCCCTCTGTCTCGAT 1008
Oy 372 catgcttcttcgctgcacctgagcttcgacccctcggaggaggt 372
Db 1007 CTGCTTTGACATGTGATGGTGTGCGAAGACTCAACCTCTAGGGTGAT 957

US-08-801-028-92/c
Sequence 92, Application US/08801028
Patent No. 6018102
GENERAL INFORMATION:
APPLICANT: JOAN BARBARINO
APPLICANT: JESSE M. JAYNES
APPLICANT: WILLIAM BELKNAP
TITLE OF INVENTION: UBQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRO
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: STEVEN J. HULTQUIST
ADDRESS: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
STREET: 200 PARK DRIVE, SUITE 210
CITY: P.O. BOX 14329
STATE: RESEARCH TRIANGLE PARK
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: APPLE MACINTOSH

